GENETIC EPIDEMIOLOGY OF ALLERGY TO BETALACTAM ANTIBIOTICS

Badin, Sevil

Awarding institution:
King's College London

GENETIC EPIDEMIOLOGY OF ALLERGY TO BETA-LACTAM ANTIBIOTICS

Sevil Badin

This dissertation is submitted for the degree of Doctor of Philosophy

September 2016

Declaration

I herby declare that this dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration except where specifically indicated in the text and acknowledgment. No parts of the work described in this thesis have been published yet. Parts of the work described in this thesis have been presented in an oral and poster presentation to the British Society for Allergy and Clinical Immunology annual meeting 2013. The work presented in this thesis has not been previously submitted for a degree, diploma or other qualifications at any other university and is not being concurrently submitted for any degree.

This thesis was undertaken in the Department of Medical and Molecular Genetics, the Department of Twin Research and Genetic Epidemiology and the Department of Respiratory Medicine and Allergy under the supervision of Dr. Michael Weale, Dr. M. Rosario Caballero, Dr. Kourosh Ahmadi (primary supervisor for first 18 months of PhD), and Dr. Massimo Mangino (tertiary supervisor). This dissertation does not exceed 60,000 words as stipulated by the Faculty of Life Sciences and Medicine.

Signed:

Date:

FOR MY BELOVED PARENTS AND BROTHER

Acknowledgements

This thesis would not have been possible without the generous support and encouragement from my supervisors Dr. Michael Weale and Dr. M Rosario Caballero. I owe my deepest gratitude to them for their caring supervision, their enthusiastic involvement in these projects, and their invaluable guidance and advices. Their optimism and endless energy has undoubtedly changed my view to the field of genetic molecular medicine and pharmacogenomics research and impacted my career ambitions. I will be indebted to them for the rest of my academic life.

I would also like to acknowledge the support and advice I received from Dr Kourosh Ahmadi, Dr. Massimo Mangino and Professor Tim Spector during these years. They had invaluable input and support for different projects of this thesis. Dr. Michael Weale has been involved in most of the statistical works of the thesis. I hope that my collaboration with him will continue for several coming years. I would also like to thank Dr. Michael Weale and Dr Alireza Moayyeri for their kind advice, which led to the methodology for estimation of heritability, GWAS and MWAS in my thesis.

A special thank you should go to the Department of Medical and Molecular Genetics and the Department of Twin Research and Genetic Epidemiology at King's College London and the Division of Allergy at Guy's Hospital, who were directly or indirectly involved in this thesis. In particular, I would like to thank Dr. Adaikalavan Ramasamy and Idel Erte for their support and enthusiasm in sharing their vast knowledge and skills. I would also acknowledge the support

and friendship from Dr. Afrooz Kosari and Dr. Elham Hasanzadeh. The support from my family members has been invaluable during this journey, including all my uncles; Hamo, Naso, Aziz, Jalal and Saied. My beloved cousins; Nojan, Arshia, Dena, Rojia, Nikia, Rudabeh and Roham. I would like to specially thank my beloved aunties; Sholeh, Leila, Shahla, Farrin and Donya.

My PhD journey was a lovely experience and I shared this with several PhD students in London. I am also grateful to all my friends who made life in London a lovely experience for me.

Lastly, and most importantly, I am truly and whole-heartedly thankful to my parents Abdollah Badin, Jaleh Hatami and brother Sam Badin for their unconditional love and support, without whom none of these achievements would have been even imaginable. They have always supported and encouraged me to be the best in all matters of life. To them I dedicate this thesis.

Abstract

Background: Immediate antibiotic allergic response is an important public health problem. Genetic and molecular characterization will improve treatment outcomes for truly allergic patients, and also reduce the use (and risk of evolution of pathogen resistance) of second-line antibiotics given to patients who incorrectly believe themselves to be allergic to first-line antibiotics.

Objective: To identify genetic and metabolic factors associated with allergic responses to beta-lactam antibiotics using the TwinsUK cohort and recruited participants from the Guy's allergy clinic.

Methods: The TwinsUK cohort is the largest registry of adult twins in the UK, and the Guy's allergy clinic is an outsized clinic covering a large area of the UK. The TwinsUK cohort has been extensively molecularly characterized. After characterising the heritability, we conducted the first high-coverage genome-wide association study (GWAS) between 211 self-reported cases in the TwinsUK cohort with questionnaire-defined beta-lactam allergic status and over 1000 individuals without self-reported allergic reaction to any substances. Approximately 2.1 million imputed and genotyped single nucleotide polymorphisms were investigated. A second GWAS was conducted on 48 clinically proven cases from the Guy's Hospital allergy clinic and ~6000 population controls. In addition a metabolome-wide association study (MWAS) was conducted on the same TwinsUK registry individuals, scanning 510 different metabolites.

Results: Following refinement of the self-reported beta-lactam allergy phenotype via the application of a more detailed questionnaire, we estimated a heritability of 21%. The heritability estimates provided positive evidence for a genetic component for beta-lactam allergy. A single hit from the TwinsUK GWAS at the MTHFS/BCL2A1 locus was found ($p<5x10^{-8}$), indicating a provisional "genome-wide significant" hit. Results from the TwinsUK MWAS demonstrated that all metabolites responded as a correlated system to the differences among twins in their allergy status. There were also 4 distinct "metabolome-wide significant" hits, of which two corresponded to known metabolites, suggesting that people who had penicillin allergy had less piperine in their system in comparison with our control group and had higher amounts of 4-vinylphenol-sulfate metabolite.

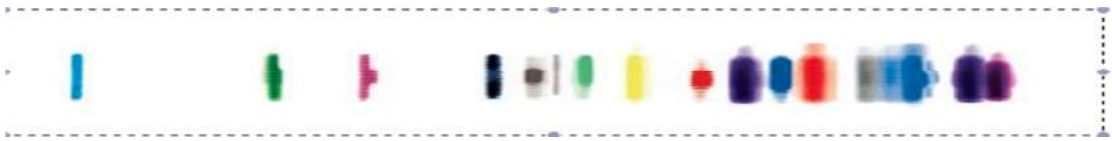
Conclusion: This study demonstrated a genetic component to beta-lactam allergy, and in particular provided evidence for a genetic signal at the MTFHS/BCL2A1 locus. Although the MWAS study showed that there was a metabolomic difference between the allergic and non-allergic individuals. These findings may lead to new personalised treatments based on a combination of genotyping and metabolic characterization. The findings of our studies need verification in independent cohorts.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction



1.1 Allergy and hypersensitivity

Hypersensitivity reactions are described as harmful immune reactions that occur in response to inherently harmless environmental antigens such as food, pollen and drugs. Allergy is the most prevalent type of hypersensitivity reaction. The term allergy is defined as an abnormal adaptive immune response of the body against non-infectious and harmless environmental substances. The facilitation of allergy can be attributed to either cell- or antibody- mediation, the latter being typically responsible for allergic reactions. Johansson et al (2001) proposed a revised nomenclature for allergic diseases to encourage uniform description, as this is a complex area where often more than one type of allergic disease presentation is found in a particular individual. The organ-based classification schema includes: rhinitis, conjunctivitis, asthma, skin diseases, urticaria and contact eczema/dermatitis – all of which are broad categories which will be fully explained in the content of this thesis and, in particular, their relatedness to allergic reactions as they are not allergic specific symptoms. The allergen source can also be used to classify allergic reactions such as food, drug and venom allergy (Baldo and Tovey, Johansson et al., 2001). Hay fever, asthma, hives, eczema, insects, drugs and food allergies are amongst the most common allergic reactions (Cookson, 1999).

In recent years, empirical studies conducted in many countries have reported a significant increase in the frequency of allergic reactions and of hypersensitive allergic reactions. The most rapid increase in prevalence of atopic diseases has occurred in Western countries in the years between 1960 and 1990, and various hypotheses have been proposed to explain this increase in prevalence. These

include: increased awareness and improved diagnostics; genetic susceptibility; psycho-social influences; increased allergen exposure; and increased environmental pollution. Interestingly, significant differences in the prevalence of symptoms of different disease have also been found between countries. The highest asthma prevalence rates have been found in Australia, New Zealand, the United Kingdom, the USA and some cities in Latin America. Lower prevalence rates have been found in non-industrialized countries and more-rural areas (Ring et al., 2001, Wold, 1998).

1.1.1 Classifications of hypersensitivity reactions

In 1963, Phillip Gell and Robin Coombs proposed a seminal formal classification system for hypersensitivity reactions to foreign material (Gell and Coombs, 1963). The classification is known as the Gell and Coombs classification, and the hypersensitivity reactions were classified based on immune mechanisms (Descotes and Choquet-Kastylevsky, 2001a). As shown in Table 1.1, all hypersensitivity reactions were divided into four different types, providing clarity to a clinically imperative field (Gell and Coombs, 1963).

Table 1.1 Gell and Coombs' classification of hypersensitivity reaction. Based on information from (Descotes and Choquet-Kastylevsky, 2001a).

Type	Other alternative names	Mediators	Example of hypersensitivity reactions
I	Immediate allergy	IgE	Atopy, Anaphylaxis, Asthma
II	Antibody-dependent, cytotoxic	IgM or IgG	Some drug allergies (e.g. penicillin)
III	Complex immune disease	IgG	Serum sickness
IV	Delayed reaction	T-cell	Contact dermatitis

According to Gell and Coombs' classifications, hypersensitivity reactions were classified into four types I, II, III and IV. This table describes each type of reaction, underlying mechanisms, any other alternative and common names for the reaction, and an example of hypersensitivity reactions for each type.

1.1.2 Mechanisms of hypersensitivity reactions according to Gell and Coombs' classification

Type I

Type I hypersensitivity reactions are known as immediate hypersensitivity reactions. These are mediated by immunoglobulin E (IgE) antibodies to a specific antigen. The primary cellular component in immediate hypersensitivity reactions is the mast cell or basophil. During sensitisation, the cell binding nature of IgE to high affinity receptors on the surface of mast cells and basophils promotes degranulation of the cells, which results in a rapid and massive release of mediators such as histamines. The symptomatic clinical presentation of type I IgE mediated hypersensitivity reactions vary in severity, and include rhinitis,

urticaria, angioedema, and potentially fatal anaphylactic shock. These reactions can occur within a few minutes to an hour of exposure (Rajan, 2003b).

Type II

Type II hypersensitivity reactions are known as cytotoxic reactions, mediated by immunoglobulin M (IgM) or immunoglobulin G (IgG) antibodies. The pathophysiological mechanism of cytotoxic reactions is described in more detail later in this chapter. A special category of type II responses involves IgG antibodies against cell-surface receptors that disrupt the normal function of the receptor (Descotes and Choquet-Kastylevsky, 2001b).

Type III

Type III hypersensitivity reactions are known as immune complex-mediated reactions and may involve tissue injury by immune complexes. Type III hypersensitivity reactions occur when antigens react with antibodies and form antigen-antibody complexes in the blood, and thereby cause secondary damage to cells (Rajan, 2003b).

Type IV

Type IV hypersensitivity reactions are known as T cell-mediated, delayed reactions. The symptoms usually develop within 2-14 days after exposure to the allergen (Descotes and Choquet-Kastylevsky, 2001a).

We should note that these hypersensitivity reactions (types I-IV) are tightly connected. As an example of these connections, we can note that maturation of B cells to IgE- or IgG-producing plasma cells all depend on the help of T-cells (Pichler, 2007b).

In 2001, Descotes and Choquet-Kastylevsky discussed whether Gell and Combs' classification remained valid (Descotes and Choquet-Kastylevsky, 2001a). They listed a number of limitations of the classification, such as the diversity of immune-mediated hypersensitivity reactions to drugs being far beyond the limited categorisation by Gell and Coombs, including such reactions as toxic epidermal necrolysis (Lyell's syndrome) and 'immuno-allergic' hepatitis. Also, allergic drug reactions can be produced by more than one mechanism, and in this case the reaction cannot be categorised by this simple classification. For example most of the beta-lactam antibiotics can cause anaphylaxis, serum sickness or immuno-allergic haemolytic anaemia, which cannot be pigeonholed by this classification.

One of the main arguments in the article was that this classification should no longer be used as a general basis for implementing strategies in preclinical drug evaluation, because of it not taking into account the complexity of dose dependent reactions (Descotes and Choquet-Kastylevsky, 2001a). Another limitation of the Gell and Coombs classification is that the classification was developed before the discovery of the functional heterogeneity of T-cells (Demoly and Hillaire-Buys, 2004). Nevertheless, it remains a simple and logical classification with mechanism-based classification to distinguish immune sensitivities (Descotes and Choquet-Kastylevsky, 2001b, Rajan, 2003b).

1.2 Aetiology of allergy

Allergy is a complex trait, which is believed to be the result of variation within multiple genes and their interaction with behavioural and environmental factors (De Swert, 1999, Kaiser, 2004). In developed countries, environmental factors and lifestyle changes are becoming more significant risk factors for allergic reactions. Also, molecular studies have identified some specific genetic variations as risk factors for allergic reactions (Kaiser, 2004). In other words, both genetic and environmental factors are involved in allergic reactions. Risk factors for allergy can be divided into two general groups: host and environment. Table 1.2 contains examples of host and environmental risk factors for allergic reactions. Gender, race, age and genetics are examples of host related risk factors for allergic reactions. Environmental pollution, exposure to infectious agents and dietary changes are some of the examples given for environmental risk factors for allergic reactions (Kaiser, 2004).

Table 1.2 Risk factors in allergic reactions.

Host	Genetics, race, gender, age
Environmental	Environmental pollution, dietary changes, exposure to infectious disease

This table shows the host and environmental factors causing allergic reactions. Genetic variations, race, age and gender are examples for host related risk factors. Environmental pollution, dietary changes and exposure to infectious disease are some of the examples of environmental risk factors of allergic reactions. Based on information from Kaiser (2004).

Host factors and their effects on severity of allergy are the major themes of this dissertation, and hence will be discussed in detail in the forthcoming chapters. However as mentioned above, allergic reactions can also occur due to environmental factors, through exposure to harmless and non-infectious environmental substances (Zuberbier et al., 2006).

The factors involved in allergic reactions are not yet all fully understood. This has caused difficulties in understanding the biological pathway and the causes of the recent increase in the prevalence of allergies.

1.3 Drug hypersensitivity

Allergic reactions are one of the major factors that limit the use of drugs and affect the therapeutic benefits of medicines. Demoly *et al.* (1999) define drug allergy as a hypersensitivity reaction where immunological mechanisms, namely IgE or T-cells, are involved.

1.3.1 Epidemiology of adverse drug reactions

The response to medication is very variable and unpredictable, and sometimes can be fatal. In developed countries, up to 5% of all hospital admissions are reported to be due to adverse drug reactions (ADR) (Pirmohamed et al., 1998). ADRs are estimated to cost the National Health Service (NHS) approximately

£466 million per year (Pirmohamed et al., 2004). Causes of variability in drug response are complex and may be due to various factors such as age, sex, pre-existing diseases, environmental and genetic factors (Nguyen et al., 2006).

1.3.2 Adverse drug reactions

The World Health Organization (Smith et al., 2013) describes Adverse Drug Reactions (ADR) as any unintended, undesired, and noxious effects of drugs which occur at doses used for prevention, diagnosis or treatment (Khan and Solensky, 2010).

Adverse drug reactions are divided into two main types, type A reactions (pharmacological) and type B reactions (idiosyncratic) (Pirmohamed et al., 1998).

Table 1.3 shows the classification of adverse drug reactions. Type A reactions are reported as being predictable in 80% of the cases. Type B reactions, on the other hand, are not predictable. Type A reactions are divided into 3 categories, according to pharmacological adverse effects of the drug, drug interactions and others. Type B reactions are divided into non-immune mediated and immune-mediated reactions. Most of the type B reactions are immune-mediated and can be partly explained by the Gell and Coombs classification (Schnyder and Pichler, 2009).

Table 1.3 Classification of adverse drug reactions.

Type A (Predictable)	Pharmacological adverse effects Drug interactions Other
Type B (Not predictable)	Non-immune-mediated Immune-mediated

Table shows the classification of adverse drug reactions. ADRs are divided into two types. Type A reactions are predictable and type B reactions are not predictable and not dose dependent. Based on information from (Schnyder and Pichler, 2009).

Type A reaction

Type A adverse drug reactions are augmented pharmacologic effects of a drug. These types of reactions are usually dose-dependent. Dose-dependent ADRs are known to be reversible either by reducing the drug dose or withdrawing the drug. Between the two types of ADRs, type A are more common and they have been reported as being responsible for over 80% of the adverse drug reactions (Pirmohamed et al., 1998).

Type B reaction

The term idiosyncratic refers to something that is specific to an individual. Therefore in the case of type B or idiosyncratic adverse drug reactions (IDRs), it is not possible to predict who will develop the reaction to a specific drug (Uetrecht and Naisbitt, 2013).

Type B adverse drug reactions are less common, but they are often life threatening. The mechanisms of IDRs are not involved in the therapeutic effect of the drug and individual susceptibility is determined by host specific factors. (Pirmohamed and Park, 2003).

1.3.3 Clinical characteristics of IDRs

The underlying mechanisms of IDRs are still unclear, but clinical characteristics of the reactions provide mechanistic clues (Uetrecht and Naisbitt, 2013). Type B (idiosyncratic adverse drug reactions) can affect different organs such as liver, skin, heart and muscle and skin (Daly, 2013). Similar pattern of IDRs can be caused by different drugs, but there are clear common clinical characteristics for most IDRs (Daly, 2013).

1.3.3.1 Skin rash

Skin rash is known as the most common type of IDR. This can be explained by the high immunological activity of the skin and visibility of even a very mild reaction on the skin (Uetrecht and Naisbitt, 2013). Below is a list of different skin rashes caused by type B ADRs, listed from most common to least common.

1. Maculopapular rash.
2. Urticaria

3. Fixed drug eruption
4. Drug-induced hypersensitivity syndrome
5. Acute generalized exanthematous pustulosis
6. Steven-Johnson syndrome and toxic epidermal necrolysis

1.3.3.2 Liver injury

There is heterogeneity in the phenotype of the adverse drug reactions affecting the liver. All the idiosyncratic adverse drug reactions affecting the liver are referred to as IDILI. Idiosyncratic liver injury (IDILI) is the most common liver injury caused by IDRs. The liver is the major site of drug metabolism and IDILI is known as one of the most common reasons of drug withdrawal (Daly, 2013). Below is a list of some of the different liver injuries caused by IDRs, listed from most common to least common.

1. Hepatocellular
2. Cholestatic
3. Hepatocellular/cholestatic mixed
4. Other types such as fibrosis induced by methotrexate

1.3.3.3 Hematologic adverse reactions

IDRs can also affect blood cells by disturbing production of blood cells or damaging blood cells (Uetrecht and Naisbitt, 2013). Below is a list of some of the different IDRs involving blood cells.

1. Agranulocytosis
2. Thrombocytopenia
3. Anemia
4. Aplastic anemia

1.3.3.4 Drug-induced autoimmunity

There is a long list of drugs that can be involved in causing autoimmune syndromes. Most of the drug-induced autoimmunity responses are resolved when the drug is stopped (Uetrecht and Naisbitt, 2013). Below is a list of some of the different autoimmune syndromes caused with IDRs.

1. Drug-induced lupus like syndromes
2. Drug-induced cutaneous lupus
3. Organ specific autoimmunity

1.4 Drug hypersensitivity reactions

Pirmohamed et al. (2003) considered drug hypersensitivity reactions to be a classic example of type B adverse drug reactions. They explained that hypersensitivity reactions to drugs are not predictable; and also that they are not dose dependent. As with type B ADRs, drug hypersensitivity reactions affect a minority of the patients, which likely means host-related factors are involved in the reactions. Also hypersensitivity reactions cannot be replicated in animal models (Pirmohamed and Park, 2003).

Variation in the incidence of drug hypersensitivity reactions can be explained by the type of drug, ethnicity of the patients and underlying disease which is being treated by the drug (Pirmohamed, 2006). The frequency of reporting of hypersensitivity reactions with mild symptoms such as skin rashes for patients undergoing active treatment is higher than that for cases with serious hypersensitivity syndromes. Skin rash is reported in up to 16% of the patients, and hypersensitivity reactions occur in 1 in 5000 to 10000 patients treated with particular drugs (Pirmohamed, 2006).

1.4.1 Types of drug hypersensitivity reactions

Determining the type of the hypersensitivity reactions is mainly based on the involved immune response and the site of the antigen formation during the reaction (Pirmohamed et al., 1998). As previously discussed, drug hypersensitivity reactions are classed as type B adverse drug reactions, which

are immune-mediated and not predictable. But it is important to underline that some of these reactions are reported to be highly associated with expression of specific human leukocyte antigen (HLA) alleles (Pichler et al., 2011). I will briefly explain some of these predictable reactions later on. Immune-mediated type B reactions can be classified into 4 different types.

Type I: IgE-mediated drug hypersensitivity

Type I reactions are IgE mediated and the mechanism of this type of hypersensitivity reaction is well understood. Hypersensitivity reactions that are induced by penicillins are a well-known example of type I reaction. The Incidence of severe IgE mediated hypersensitivity reactions is reported to be 1 in 2000 patients (Pirmohamed et al., 1998).

Primary drug sensitization can cause the development of drug specific IgE. After the formation of specific IgE in the body, even the reintroduction of a small amount of the drug (antigen) can induce symptoms. In type I (IgE-mediated) reactions, the allergen (in this case drug) must bind to antigen-binding receptors of the corresponding IgE molecules. Type I reaction is also known as an immediate reaction and usually occurs within 1 hour after drug administration (Schnyder and Pichler, 2009).

Type II: IgG-mediated cytotoxicity

Type II (IgG-mediated) reactions are cytotoxicity reactions directed to the membranes of cells in the bone marrow, such as erythrocytes, leukocytes, platelets and perhaps hematopoietic precursors. Drugs such as methyldopa and

aminopyrine are examples of drugs inducing type II reactions (Schnyder and Pichler, 2009).

Type III: immune complex deposition

Type III reactions result from immune complex deposition. Immune complexes are commonly formed during normal immune responses in the body. These reactions normally occur without any clinical symptoms, and except in some rare cases they are formed as immune complexes bound to endothelial cells. The attachment of immune complexes to endothelial cells results in immune complex deposition with complement activation in small blood vessels. Symptoms such as serum sickness, vasculitis and lupus erythematosus are reported for type III reactions. Serum sickness is the clinical symptom of beta-lactam induced type III reaction. Also, drug-induced lupus erythematosus is a clinical symptom of type III reactions induced by quinidine (Schnyder and Pichler, 2009).

Type IV: T-cell-mediated drug hypersensitivity

The underlying mechanisms of T-cell mediated drug hypersensitivity reactions are not yet fully understood. This type of hypersensitivity reaction is cell-mediated. Type IV reactions are also known as delayed reactions. Cellular response is generally visible 48-72 hours after drug exposure. One hypothesis is that type IV hypersensitivity reactions may possibly be explained by the hapten/prohapten and the p-i concepts, as described below (Pirmohamed et al., 1998).

1.5 Hapten/prohapten concept

Components with low molecular weight (<1000D), such as drugs, are too small to be able to cause an immune response in the body. The hapten/prohapten concept is a proposed mechanism to explain how these small molecules can stimulate an immune response. The hapten concept also explains how chemically reactive drug or drug metabolites can bind covalently to proteins and modify them, and how the modified protein can induce an immune response (Pichler, 2008). Haptens are chemically reactive small compounds that bind to proteins or peptides and modify them. There are two main mechanisms by which haptens can be responsible for inducing an immune response (Pichler, 2008, Pichler et al., 2011).

1. **Stimulating the innate immune system**

   Covalent binding of hapten to cellular protein can stimulate the innate immune system. The covalent binding of hapten to cellular protein can transmit a danger signal, and the danger signal can be the start of the innate immune system stimulation (Pichler, 2008).

2. **Stimulating the specific immune system**

   Haptens can stimulate the specific immune system by forming hapten-protein complexes. The hapten-carrier complex is presented as a hapten-modified peptide to T-cells. Hapten modification can happen with cell-bound proteins, soluble proteins or with the major histocompatibility complex (MHC)(Pichler, 2008).

1.5.1 Prohaptens

As with haptens, prohaptens are really small molecules that cannot cause any immune response by itself. Prohaptens are not chemically reactive, and therefore they are not able to bind to proteins via a covalent bridge. These molecules need to be converted into a hapten form by being metabolised in order to become chemically reactive (Pichler et al., 2011). Prohapten metabolism mainly occurs in the liver, which is usually considered to be a tolerogenic environment and thus the metabolised hapten may not cause an immune response there (Pichler et al., 2011).

1.5.2 Pharmacological Interaction with immune receptors (p-i concept)

As explained above, the drug or the metabolites of the drug need to be chemically active in order to covalently bind to the proteins and modify them to induce an immune response. But what is missing here is how the immune response occurs when the drug is lacking hapten characteristics. In order to explain the underlying mechanisms of immune response to these drugs, the hapten/prohapten concept has been supplemented by the p-i concept. The p-i concept stands for direct 'pharmacological interaction' of drugs with immune receptors (Pichler, 2008).

The p-i concept proposes that drugs can bind directly and reversibly to immune receptors and stimulate the cells to cause an immune response in the body. The p-i concept can be used in order to explain the hypersensitivity reactions caused by drugs that are antigenic in their parental form; drugs that are not in reactive form; drugs that do not require antigen processing; and finally drugs whose antigen presentations involve reversible binding (Pichler et al., 2011).

The large variability in unpredictable drug response is a big health concern. The main aim of clinicians when they prescribe a drug is to treat patients and cure the underlying disease or condition without causing any adverse drug reactions in the patient. Prescribing a particular drug to a patient with a particular disease needs a careful assessment of the benefit-risk relationship. Pirmohamed (2014) highlighted several issues in drug selection and dose recommendation by prescribers (Pirmohamed, 2014).

One limitation refers back to the drug development phase. During the drug development process, most of the time the drug dosage is defined based on mean values, and often other parameters such as body surface area are not taken into account. Therefore clinicians implicitly assume that all patients will respond to the drug in the same way.

A second limitation is our ignorance of the factors that can affect the pharmacokinetics of the drug (pharmacokinetics is the study of what the body does to drugs). Patients vary according to factors such as ethnicity, age, weight, smoking status, dietary habits and alcohol intake. Lack of knowledge regarding interacting medications is another factor in the list of ignored factors.

A third limitation is that diseases are classified according to easily measurable phenotypic markers, but there may be disease subphenotypes which respond differently to drug treatment. Developing better molecular markers to stratify drug treatments is still needed to personalise treatment.

Figure 1.1 shows the elements required to achieve the goal of personalised medicine and how the factors discussed above are important in this process (Pirmohamed, 2014).

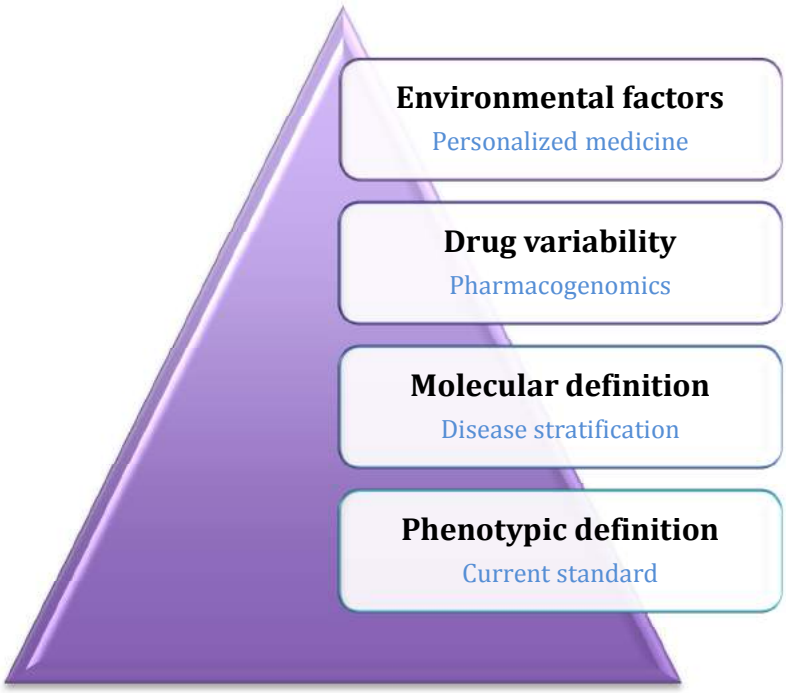


Figure 1.1 Necessary required elements to accomplish personalised medicine. Based on information from (Pirmohamed, 2014).
This figure gives information on the different elements that are necessary in order to achieve the goal of personalised medicine.

1.6 Pharmacogenomics

Pharmacogenomics is the study of genetic factors affecting drug response. The term pharmacogenomics may also be expanded to the development of new drugs through the use of genomic information (Daly, 2010). After 50 years of studying the genetic basis of variation in drug responses in individuals, it is now clear that

there are many polymorphisms in human genes that can have a significant functional effect (Daly, 2010). The incidence of more than one allele at the same locus with appreciable frequency among the same population is defined as a genetic polymorphism (Daly, 2010).

Identifying the relationship between response variables and genetics will possibly be used to identify people with good response, and those who will suffer adverse drug reactions, early. This will lead to personalised medication, which is supported by genetic analysis and which aims to achieve the optimal medical outcomes for each individual. The aim is to improve the quality of life, health and potentially reduce overall health-care cost (Altman et al., 2011).

Kroemer and Meyer zu Schwabedissen (2010) define personalised medication as the management of a patient's disease or disease predisposition. From the clinical point of view, pharmacogenomics should increase the efficiency of the prescribed drug and reduce the morbidity associated with adverse drug reactions. From a pharmaceutical industry viewpoint, pharmacogenomics should help to reduce costs and timelines in clinical trials (Kobayashi and Satoh, 2009).

In the future, drugs may be developed based on biological markers (DNA, RNA and proteins) of each individual. These drugs will, it is hoped, maximise therapeutic effects and minimise damage to healthy cells. Therefore, there will be no need for the conventional trial-and-error method to finding the right drug that matches each patient's medical requirement. To this end, analysing a patient's genetic profile will provide an opportunity to prescribe the best

available drug for individuals from an early stage of diagnosis. This will also reduce the time of recovery and the probability of any adverse drug reactions.

Another benefit of pharmacogenomics is that it will provide a more accurate method to discover how an individual's body processes medicines, and how long it takes to metabolise a drug. This will also help to determine the appropriate drug dosage. Furthermore, in the case of complex diseases, having the knowledge of being susceptible to certain disease will allow a person to change his/her lifestyle and environment in the early stages of life. It also offers the advantage of early monitoring and treatment, which can be introduced at a more efficacious stage (Altman et al., 2011).

1.6.1 Pharmacogenomic approaches to identify genes that contribute to adverse reactions

In the early years, pharmacogenetic investigations of drug response focussed on the function of a single gene. However, in recent times, studies have shown that most drug responses occur as a result of a complex interaction between genes and environment factors (Pirmohamed, 2006), motivating other approaches.

Pharmacogenomic studies of adverse drug reactions typically involve case/control association studies to investigate genetic susceptibility for the reaction. Case/control studies are either done on a candidate gene approach (just looking for the specific region of the gene) or as a genome-wide association

study (GWAS), which covers the whole genome (Daly, 2013). More detail on the genome-wide approach is provided in chapter 3.

1.6.2 Examples of pharmacogenomic studies

The major histocompatibility complex (MHC) is a set of cell surface molecules encoded by a large gene family called human leukocyte antigens (HLA), situated on chromosome 6 (Shiina et al., 2009). Given the role of the MHC in immune response, many pharmacogenetic studies have focused on this region, and both candidate gene and genome-wide association studies on different adverse drug reactions have reported significant associations with HLA genes (Daly, 2013). A detailed understanding of the association between HLA alleles and drug hypersensitivity provides the benefit that these severe and unpredicted ADRs might one day become predictable. Here I shall give some examples of drug responses associated with the HLA region.

1.6.2.1 HLA association with DILI

DILI is a rare form of idiosyncratic adverse drug reaction associated with commonly used drugs such as flucloxacillin. DILI is known to be a major cause of post marketing drug withdrawals. The incidence of flucloxacillin induced DILI has been reported as 8.5 in every 100,000 patients treated with flucloxacillin in the UK. Flucloxacillin is one of the widely used antibiotics for infections caused

by staphylococcal bacteria in some European countries (Daly et al., 2009). The pathogenesis of DILI generally involves the participation of the toxic drug or its metabolites that stimulate the body's immune response or directly affect the cell. These reactions can cause cell death and lead to hepatotoxicity. As discussed before, metabolism of drugs largely takes place in the liver, thus the liver is one of the main organs that is affected by metabolism dependent drug induced injury (Daly et al., 2009).

A genome-wide association was conducted on 51 cases with flucloxacillin induced liver injury 282 controls matched for sex and ancestry in UK in 2009 (Daly et al., 2009). This genome-wide association study showed strong association for a SNP (rs2395029), with a p-value of 8.7 □x 10□$^{-33}$ in the MHC region with flucloxacillin induced liver injury. The SNP (single nucleotide polymorphism) was in complete linkage disequilibrium (LD) with *HLA B*5701*. The novel finding of the study was then replicated in another cohort with 23 cases.

Another genome-wide association study reported a significant association between amoxicillin-clavulanate induced liver injury and class I and class II HLA genotypes (Lucena et al., 2011). Amoxicillin-clavulanate (AC) is one of the most commonly prescribed antimicrobial agents worldwide, and the clavulanate component is known to be the main cause of AC induced DILI. This study conducted a genome-wide association study on 201 cases with DILI and 532 population controls. They reported many associated loci in the MHC region for AC induced DILI, but the most strongly associated SNP (rs9274407, p= 4.8 x 10^{-}

[14]) was in LD with HLA-DRB1*15:01 (a class II MHC molecule). This SNP was correlated with the previously reported SNPs that associated with AC-DILI. They also reported a new SNP in LD with HLA-A*02:01 (a class I MHC molecule).

Genetic susceptibility is a fundamental component of serious adverse drug reactions. There is a growing interest in the possibility that developing genetic tests to identify all those at risk of adverse events prior to prescription in a screening programme may lead to valuable drugs being retained. Thus, I now discuss examples of three genetic association studies on abacavir, carbamazepine and warfarin induced hypersensitivity reactions.

1.6.2.2 HLA association with carbamazepine induced hypersensitivity reactions

An association between carbamazepine-induced skin rash and HLA B*15:01 was first reported in a Taiwanese study, using a candidate gene approach (Chung, 2004). Because is very rare in non-Han Chinese, studies on Europeans and Japanese failed to show that this HLA allele was a risk factor for carbamazepine skin rash in these populations. McCormack et al, (2011) presented a GWAS which aimed to detect novel genetic risk factors for carbamazepine-induced skin rash in a European population. McCormack et al. reported a novel HLA association involving HLA-A*31:01 but, as expected, did not find an association with HLA B*15:02. The HLA-A*31:01 allele was significantly associated with the

hypersensitivity syndrome in the Northern European sample, and maculopapular exanthema was also found to be associated with HLA-A*31:01.

1.6.2.3 HLA association in abacavir induced hypersensitivity reactions

The association of HLA-B*57:01 with abacavir was first reported using a candidate gene approach in 2002 (Mallal, 2002). A follow-up study investigated the merits of a screening test based on this association, via a double-blind randomized control trail (N=1956) conducted in 19 different countries (Mallal et al., 2008). This latter study is an important example of translational research in pharmacogenetics, translating the important findings of allele associations to hypersensitivity reactions and extending this knowledge to provide screening opportunities in clinical practice settings to help improve patient health outcomes. The authors found that genetic screening eliminated all cases of immunologically confirmed hypersensitivity reaction in their study (0% in the screened group vs. 2.7% in the control group, P<0.0001), with a negative predictive value of 100% and a positive predictive value of 47.9%) (Mallal et al., 2008). This study demonstrated a clear benefit for genetic screening and this led to the widespread implementation of testing before prescribing abacavir.

1.7 Allergy to beta-lactam antibiotics

Beta-lactam antibiotics are widely prescribed for the treatment of common infections caused by a range of bacteria. They are also used as first-line prophylaxis in surgical procedures (Mirakian et al., 2015). Penicillins and cephalosporins are the two major classes of the beta-lactam family (Pichler, 2007a). In the penicillin category, amoxicillin is the most prescribed beta-lactam antibiotic in Europe (Treudler and Simon, 2007). Due to their widespread use, beta-lactam antibiotics, especially penicillins and cephalosporins, are known to be the most common causes of drug hypersensitivity reactions (Lin et al., 2010, Mirakian et al., 2015).

1.7.1 Types of beta-lactam allergy reactions

Depending on the time interval between taking the drug (e.g. beta-lactam antibiotics) and the appearance of the symptoms, allergic reactions to beta-lactams are classified into two types (Mirakian et al., 2015).

The first type of reaction is called an immediate reaction, which is mediated by an IgE response, and symptoms appear within 1 hour of introducing the drug (Mirakian et al., 2015).

The second type of reaction is called a non-immediate or delayed reaction, with an interval of 1 to 48 hours. This type of reaction is T-cell mediated (Mirakian et al., 2015). Blanca *et.al* (2009) argue that there are many lines of evidence

showing that the longer the interval between introducing the drug to the body and the appearance of allergy symptoms, the less likely it is that the reaction is IgE mediated (Pichler, 2007a, Blanca et al., 2009, Treudler and Simon, 2007).

1.7.2 Epidemiology of beta-lactam allergy

Beta-lactam antibiotics are responsible for 26% of fatalities caused by drug anaphylaxis in the UK (Fitzharris, 2008). A more recent study reports that up to 20% of drug-related anaphylaxis deaths in the UK are caused by penicillin, with a higher estimate of 75% for the United States (Mirakian et al., 2015).

Up to 20% of hospitalized patients self-report beta-lactam allergy (Mirakian et al., 2015). Although there is no prospective study to show the prevalence of penicillin hypersensitivity in the general population, clinical records show that 10% of people treated with beta-lactam antibiotics have reported an allergic reaction (Demoly and Romano, 2005, Solensky, 2003). However, after conducting clinical testing only 1-10% of these patients have evidence of type I hypersensitivity reactions (Mirakian et al., 2015).

Lower prevalences for penicillin allergy have been reported in cohort studies, ranging from 0.2% per course of treatment in a large population (Mirakian et al., 2015) to between 3 and 5% in a drug surveillance study (Mirakian et al., 2015). In particular, allergic reactions to penicillin appear to be 10 fold more common compared to reactions to cephalosporins (Mirakian et al., 2015).

If patients who are wrongly labelled as allergic to the beta-lactam family are prescribed unnecessary alternative drugs, this can lead to a risk of developing multiple drug-resistant bacterial infections. These treatments can also cost more than treatment with beta-lactam antibiotics (Solensky, 2003).

In order to address the reported heterogeneity in clinical practice of beta-lactam hypersensitivity in the UK and to avoid the use of less effective and more expensive medications, a detailed guideline was provided recently (Mirakian et al., 2015).

1.7.3 Risk factors for immediate response to penicillin

Although the exact aetiology of immediate IgE-mediated allergy to beta-lactam antibiotics is unclear, it is suspected that it has a complex aetiology (Mirakian et al., 2015). Host related and environmental factors such as heritability, age, underlying infection, asthma and sex have been reported as influential risk factors in developing hypersensitivity reactions to beta-lactam antibiotics (Mirakian et al., 2015).

1.7.3.1 Genetic risk factors

The underlying genetic risk factors remain unsolved for beta-lactam allergy. However family history data show that beta-lactam induced adverse reactions

typically run in the family, and are more likely to be T-cell mediated (delayed) reactions (Mirakian et al., 2015). I will give more detail on published genetic associated studies on beta-lactam allergy later on in this chapter.

1.7.3.2 Host related risk factors

Table 1.4 provides a list of host related risk factors for penicillin allergy. A positive skin test is not necessarily obtained in patients with a positive clinical history of penicillin allergy. Therefore clinical history is not reliable on its own. In the case of childhood reactions, there is a big risk of recall bias. Also in these reactions there is a chance that the reaction was related to the underlying infection or that it was confused with drug side effects. Patients with a clinical history of having a reaction in the last 15 years have been reported to have a very low risk of a positive challenge test (0.4%) (Mirakian et al., 2015).

Women are reported to have a higher rate of penicillin allergy (11.0%) as compared to men (6.5%). The higher number of prescribed antibiotics may explain the higher frequency of penicillin reactions in women (Mirakian et al., 2015).

There is limited evidence to support the case that penicillin intake via the oral route is less likely to cause allergic reactions than via other routes such as intravenous. Topically applied penicillin is recognized as being highly immunogenic, and it is no longer used. Another host related risk factor is the frequency of administration of penicillin courses. Patients with conditions such

as cystic fibrosis often receive frequent dosages of intravenous antibiotics (Mirakian et al., 2015).

Age is known to be one of the host related risk factors for penicillin allergy. Most of the penicillin induced allergic reactions have ben reported in patients between 20 and 49 years old. But a more recent study shows that the susceptibility of having penicillin related allergic reactions increases in patients who are over 80 years old by 20% (Mirakian et al., 2015).

Table 1.4 Host related risk factors for immediate response to penicillin antibiotic. Based on information from (Mirakian et al., 2015).

Sex	Drug induced penicillin allergy has been reported more in female (11.0%)than in men (6.5%)
Frequency and route of exposure	Some evidence showing that penicillin intake via oral route is less likely to induce reactions compared to other routes
Age	Susceptibility to penicillin-induced allergic reactions increases with age
Underlying infection	HIV patients treated with co-amoxiclavand reported rashes.

The table shows information on host related risk factors associated with beta-lactam allergy. One example is given for each factor.

1.7.4 Molecular structure of beta-lactam antibiotics

There are four main groups in the beta-lactam antibiotic family, which all share a beta-lactam nuclear ring (Demoly and Romano, 2005); penicillins, cephalosporins, carbapenems and monobactams. Figure 1.2 illustrates the

chemical structure of all four beta-lactam groups, with the beta-lactam rings highlighted in red. In penicillins, apart from the beta-lactam ring that is involved in bactericidal activity, there is a five-member thiazolidine ring in common. Different penicillins are distinguished by the side chain that is fused to position 6 in the nuclear beta-lactam ring (Pichler, 2007a, Treudler and Simon, 2007, Demoly and Romano, 2005). The diversity in the side chain distinguishes the different penicillins, and it is an important factor in immunological recognition of the penicillin. Also, to some degree, the side chain is responsible for allergic cross-reactivity (Mirakian et al., 2015).

Cephalosporins have two side chains attached in position 3 and position 7. In this group, the varieties of the side chains broaden antibacterial activity by affecting drug metabolism and resistance to beta-lactamase (Mirakian et al., 2015). Carbopenems contain a beta-lactam ring and an adjacent five-member thiazolidine ring, but in the place of sulphur, there is a carbon double bond. Finally, monobactams contain only one ring (beta-lactam) and there is no fused sulphur ring (Mirakian et al., 2015).

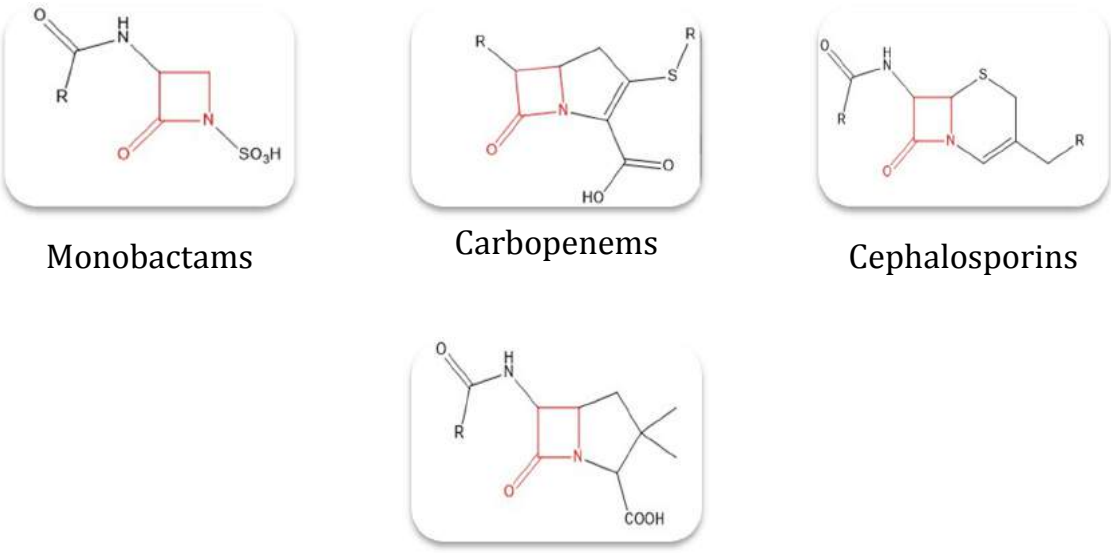


Figure 1.2 Chemical structure of the beta-lactam families. Based on information from (Torres et al., 2003).

This figure shows the chemical structure of all four beta-lactam groups. The beta-lactam ring is highlighted in red. All four groups contain a beta-lactam ring, and the main difference between these groups is the fused side chain to beta-lactam molecule.

1.7.5 Beta-lactams as an antigen

Penicillin is a chemical substance with a low molecular weight and, like many other medications, is too small to act as a complete antigen. They need to be bound to a larger carrier molecule, such as tissue or serum proteins (haptens) - a process known as haptenation (Pichler et al., 2011). The resulting new immunogenic molecule is large enough to elicit an immune response (Solensky, 2003, Rosario and Grumach, 2006).

With regard to the chemical structure of penicillins, the beta-lactam ring, the thiazolidine ring, the dihydrothiazine ring and the side chains can all be recognized as potential immunogenic molecules (Mirakian et al., 2015). Under

normal physiological conditions, the beta-lactam ring opens spontaneously and binds to the nearest protein (Solensky, 2003). Degradation of the beta-lactam ring results in a chemically unstable molecule, which can bind to lysine residue aminogroups on soluble or cell proteins. Benzyl penicilloyl is formed as a result of the beta-lactam ring binding to proteins. This newly created molecule is called the "*major determinant*", and 95% of tissue-bound penicillin exists in this form (Solensky, 2003). This molecule is known to be responsible for the majority of the beta-lactam related allergic reactions (Mirakian et al., 2015).

The remaining parts of the penicillin molecule (side chains) after binding to proteins are called "*minor determinants*" (Solensky, 2003).

1.7.6 Clinical symptoms of IgE mediated beta-lactam allergy

Symptoms of immediate allergic reactions can range from a simple local itch/wheal to generalised urticaria and edema, to severe, complex reactions called anaphylaxis. In severe cases of anaphylaxis, patients can collapse and/or become unconscious often within 15 minutes, or can even suffer a lethal anaphylactic shock (Pichler, 2007a, Pichler et al., 2010).

Clinical symptoms of immediate (IgE mediated) allergic reactions to beta-lactam antibiotics are urticarea with or without angioedema and anaphylaxis (Torres et al., 2003).

- Urticaria: wheals in one or several parts of the body;
- Angioedema: localised edema in one or several parts of the body;
- Anaphylaxis: the most severe allergic symptoms occurring immediately after introducing the drug to the body, often affecting more than one organ

Symptoms for hypersensitivity reactions are categorised into mild, moderate and severe reactions (Brown, 2004). Table 1.5 shows the grading system for seriousness of the hypersensitivity allergic reactions according to Brown (2004).

According to Brown's grading system, in mild hypersensitivity reactions, symptoms such as generalized erythema, urticaria, perorbital edema or angioedema appear in skin and subcutaneous tissues only. This grading system proposes that moderate hypersensitivity allergic reactions involve respiratory, cardiovascular or gastrointestinal effects with symptoms of dyspnea, stridor, wheeze, nausea, vomiting, dizziness (presyncope), diaphoresis, chest or throat tightness, or abdominal pain. In the case of severe reactions, symptoms include cyanosis or SpO_2<92% at any age, hypotension (systolic blood pressure (SBP) < 90 mm Hg in adults), confusion, collapse, loss of consciousness, or incontinence (Brown, 2004).

Table 1.5 Grading system for generalized hypersensitivity reactions. Based on information from (Brown, 2004).

Grade	Clinical feature	Defining symptoms and signs
1. Mild	Skin and subcutaneous tissues only	Generalized erythema, urticaria, perorbital edema or angioedema
2. Moderate	Respiratory, cardiovascular or gastrointestinal involvement	Dyspnea, stridor, wheeze, nausea, vomiting, dizziness (presyncope), diaphoresis, chest or throat tightness, or abdominal pain
3. Severe	Hypoxia, hypotension, or neurologic compromise	Cyanosis or SpO_2<92% at any age, hypotension (systolic blood pressure (SBP) < 90 mm Hg in adults), confusion, collapse, loss of consciousness, or incontinence

The table describes a grading system for the seriousness of the hypersensitivity allergic reactions. The hypersensitivity allergic reactions are divided in to 3 main categories based on the symptoms and the organs involved in the reaction.

1.8 Diagnosis of beta-lactam allergy

A detailed history of reactions and a description of the symptoms associated with the allergic reactions, as recalled by the patient, is often the first step in drug allergy diagnosis. Clinical reports are the second source of information that is useful to evaluate allergic reactions (Torres et al., 2003).

1.8.1 Clinical evaluation of beta-lactam allergy

A clear and detailed history of the drug reaction is an essential component of evaluating a patient with suspected reactions to any drugs. This clinical history should be extremely inclusive to help clinicians to choose the best possible and safest diagnostic test for the patient (Demoly and Romano, 2005, Khan and Solensky, 2010, Torres et al., 2003, Demoly et al., 2010). The clinical history should contain all the following questions:

- What was the patient's age at the time of the reaction?
- Is the patient able to recall the reaction?
- What was the name of the medication?
- What was the route of administration?
- What was the exact time of the reaction?
- What were symptoms and characteristics of the reaction?
- What was the reason for taking the medication?
- What were the other medication taken at the time?
- Did they have any prior treatment for the allergic reaction?
- Had the patient taken the same drug or any other drug with cross-reactions before/after the reaction?
- Since the reaction, has the patient used the same drug? Has the patient experienced similar reactions?

If the patient provides clear and reliable answers to all these questions, then this clinical history is deemed to be a sufficient diagnostic tool. However, the patient is not usually able to recall all details, and the doctor often cannot precisely

evaluate all the requested data (Demoly and Romano, 2005). Therefore, although a clinical history helps to distinguish between allergy and allergy-like reactions, it is not sufficient on its own to confirm the allergy (Solensky, 2003, Demoly and Romano, 2005).

1.8.2 Tests used for diagnosis of immediate allergy to beta-lactam antibiotics

The European Network for Drug Allergy (ENDA) has described a step-by-step protocol for all available tests for allergic reactions (Torres et al., 2003). However a more recent guideline on management of allergy to penicillins and other beta-lactams has been introduced to UK allergic specialists (Mirakian et al., 2015). This detailed guideline was prepared by the Standards of Care Committee of the British Society for Allergy and Clinical Immunology (BSACI) and an expert panel.

Figure 1.3 illustrates the short algorithm used for the diagnosis of immediate and non-immediate allergic reactions to beta-lactam antibiotics. This guideline is a step-by-step guide through the diagnostic process of the allergic reactions to beta-lactam antibiotics. As mentioned before, an accurate history helps in understanding the type of the reaction. This guideline highlights all the required clinical tests, all the cross-reactivity possibilities, and indicates whether the patient should be labelled as being allergic to beta-lactam antibiotics or not (Mirakian et al., 2015).

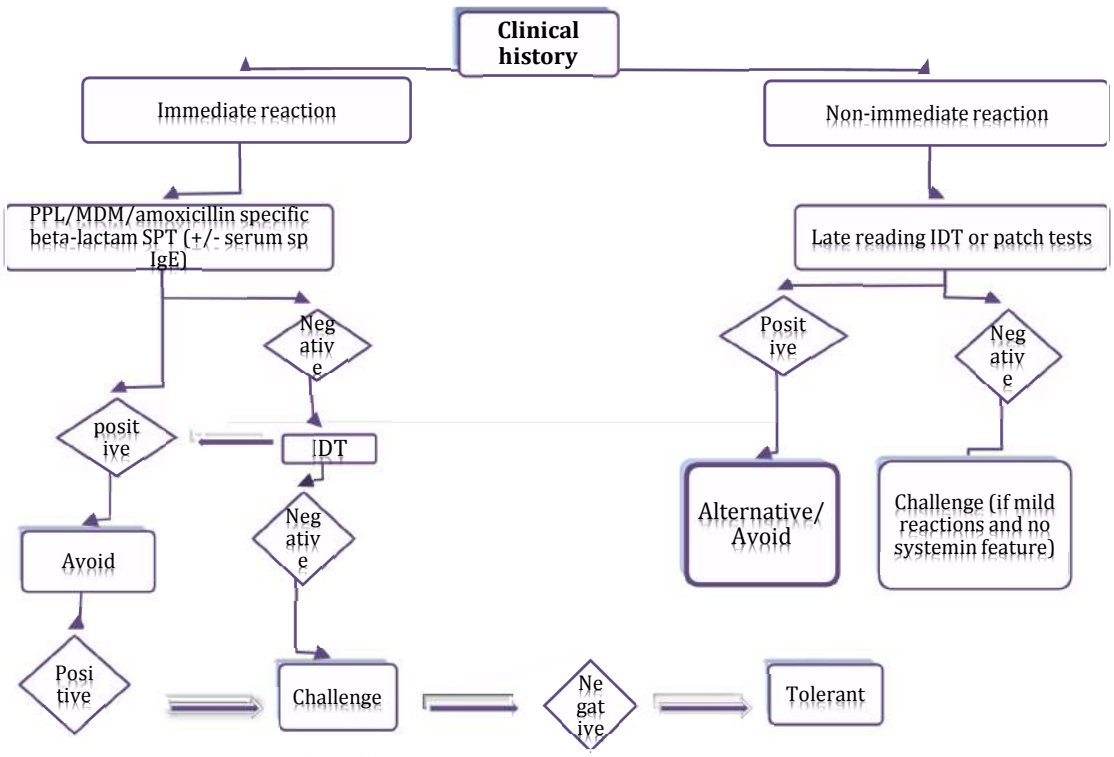


Figure 1.3 Algorithm for the diagnosis of immediate and delayed allergic reactions to beta-lactam antibiotics.
Based on information from (Mirakian et al., 2015). SPT=skin prick tests; IDT=intradermal tests.

1.8.3 Skin tests

The best known and the most widely used diagnostic tool for the diagnosis of allergic reactions to beta-lactams is the skin test (Torres et al., 2003). There are three different methods for skin tests: the skin prick test (SPT); the intra-dermal test (IDT); and the skin patch test.

In the case of beta-lactam antibiotics, skin tests usually start with a skin prick test. During an SPT, the allergen(s) is introduced to a small skin prick by the use of an appropriate needle. The process is carried out by introducing low

concentrations of the allergen, increasing to high concentrations in the following steps. During skin prick testing, patients are kept under close observation and the results are monitored after 15 minutes (Mirakian et al., 2015).

Skin tests are generally safe, but systemic reactions have been reported in 0.7-11% of those with positive skin reaction, therefore careful baseline measurements of peak expiratory flow rate, blood pressure and pulse is necessary (Mirakian et al., 2015). The results are compared with positive (histamine) and negative (diluent) controls. In the skin prick test, the appearance of a larger than 3mm wheal is considered a positive result (Mirakian et al., 2015).

According to the guideline (Figure 1.3), if the results for the SPT are negative, the clinician prepares the patient for an intra-dermal test (Mirakian et al., 2015). In an intra-dermal test, the same (hapten) allergen solution is injected into the skin (between the epidermis and dermis). Like the SPT, the results of the IDT should be evaluated after 15 minutes. A positive result from the IDT is valid if any wheal is larger than 3mm (Demoly and Romano, 2005, Kelkar and Li, 2001).

In skin tests for beta-lactam antibiotics, hapten solutions are mixtures of minor determinants (MDM) and penicilloyl-poly-L-lysine (PPL), known as the major determinant. The major antigenic determinants contain benzyl-penicilloyl, and minor antigenic determinant contains benzylpenicillin, penicilloate and penilloate (Solensky, 2003). In addition to penicillin minor and major determinants, amoxicillin and ampicillin are also considered as a part of skin

tests for beta-lactam antibiotics more generally (Solensky, 2003, Torres et al., 2003).

Between 8.4-30.7% of the patients with a negative skin test have been reported as having a positive challenge test (Messaad et al., 2004). According to Mirakian et al. (2015), patients with a positive skin test to PPL/MDM should be labelled as being allergic to all beta-lactams, and they should avoid penicillins, unless they have a negative skin test to amoxicillin/ampicillin. In this case the next step (oral challenge) is required. In the case of a negative skin test for all the determinants and amoxicillin/ampicillin, the challenge test with the original penicillin that was implicated in the allergic reaction should be reported. In the case of a negative skin test to PPL/MDM and benzylpenicillin, but a positive skin test to specific beta-lactam, the challenge test is required with penicillin V. In this case the allergy may be due to the side chain activity (Mirakian et al., 2015).

1.8.3.1 Sensitivity and specificity of skin tests

Although skin tests are known to be one of the quickest and safest ways to diagnose immediate IgE mediate allergic reactions, the specificity of the test is far less than 100% (Demoly and Romano, 2005). Toress *et al.* (2001) and Messaad *et al.* (2004) have previously reported that 8-17% of patients have shown false negative skin test results (Messaad et al., 2004, Torres et al., 2001). The sensitivity of skin tests with a positive clinical history is calculated to be 22% for PPL, 21% for MDM, 43% for amoxicillin and 33% for ampicillin.

Therefore, the overall sensitivity of skin tests is calculated to be 70% for a combination of all allergens. The specificity of the skin test is reported at 97% (Pichler, 2007a).

False-positive skin tests have also been reported. This can occur when there is no history of an allergic reaction (negative history). The incidence of false-positive skin test results is estimated at around 7% (Torres et al., 2001).

Evidence shows that a long interval between the initial reaction event and skin testing reduces the chances of a positive response. After 10 years following the allergic reaction to penicillin, only 20-30% of patients show a positive skin test. Skin test reactivity completely disappears after 5 years following an allergic reaction to amoxicillin. As a positive skin test can reduce the need for drug provocation tests, skin testing should be carried out soon after the reaction (Mirakian et al., 2015).

1.8.4 Drug provocation test

Due to the fact that there are high false negative results in the skin test, if there is a clear allergy history and a negative skin test then a drug provocation test (challenge test) is usually requested before any final clinical decisions are made (Holm and Mosbech, 2011).

The drug provocation test (DPT) is known to have the best sensitivity of all the available diagnostic tests (Torres et al., 2003, Messaad et al., 2004). However, it

is not usually requested for high-risk patients, including patients who have previously developed life threatening reactions to the drug, such as anaphylactic shock, or patients who suffer from related disease such as respiratory and cardiovascular disease (Torres et al., 2003, Holm and Mosbech, 2011).

During the drug provocation test, the drug itself is given to the patient. This process is performed under strict hospital observations requiring emergency room facilities. The process starts from a low dose of the allergen and the patient is monitored for 30 minutes. If the first dose is well tolerated, the second dose is given but with a time lapse of 30 to 60 minutes from the first dose (Torres et al., 2003, Demoly and Romano, 2005, Demoly et al., 2010).

A negative predictive value of 94% has been reported for the oral provocation test with beta-lactams. Reports show that patients with a false-negative provocation test result generally do not experience any life-threatening reactions (Mirakian et al., 2015).

1.8.5 In-vitro test

Finally, the *in-vitro* test is used as a complementary technique in the diagnosis of immediate allergy to drugs including beta-lactam antibiotics (Blanca et al., 2009). The test is based on detecting specific IgE antibodies to any beta-lactam determinants in the serum using an immunoassay. Specific IgE is detected by using the ImmunoCAP system. The ImmunoCAP capsules contain polymers which are covalently bound to the specific beta-lactam which interacts with

specific IgE in the patient's serum. Interaction can then be detected by fluorescence (Mirakian et al., 2015).

However, due to the high cost of the test and the lack of accessibility, it is not often used in diagnostic processes unless it is essential. This test is highly specific for detecting specific IgE for beta-lactams, but it is not sensitive. Also this assay is not available for most of the cephalosporins (only available for cefaclor). Sensitivity of this assay for the patients with positive skin test results to amoxicillin-and/or benzylpenicillin was reported as 54%, but, for the same patients, specificity of the same test was reported as up to 95% (Mirakian et al., 2015).

1.8.6 Misdiagnosis of immediate allergy to beta-lactam antibiotics

It is important to highlight the high incidence of self-reported penicillin allergy in the population. It is also important to bear in mind that over 90% of these "allergic" individuals are able to tolerate beta-lactam antibiotics, as has been reported in several clinical evaluations of antibiotic allergy (deShazo and Kemp, 1997, Satta et al., 2013).

Lin *et al.* (2010) tried to explain this discrepancy. They believe that there are many factors which could play a role in the over-reporting of beta-lactam antibiotic allergy, such as confusing a drug's side effects with drug allergy, or misinterpreting symptoms caused by the illness itself with allergic symptoms.

1.8.7 Cross reactivity

As explained above, all beta-lactam families share a beta-lactam ring in their chemical structure. Therefore, when the beta-lactam allergy occurs due to the beta-lactam ring, there is a cross reactivity between different beta-lactam groups (James and Gurk-Turner, 2001, Sastre et al., 1996). Recent investigations indicate that the cross-reactivity among beta-lactams is not only due to a shared beta-lactam ring but also, and mainly, due to side chain homology (Mirakian et al., 2015).

Cross-reactivity with monobactams

Monobactams are less immunogenic than penicillins or cephalosporins, due to the monobactams' chemical structure. As discussed previously, the chemical structure of the monobactam group contains a single beta-lactam ring and there is no fused sulphur ring. Aztreonam is a monobactam, and it is generally tolerated by patients who had confirmed IgE mediated or T-cell mediated reactions to beta-lactams (Mirakian et al., 2015).

Cross-reactivity with carbapenems

Cross-reactivity between penicillins and carbapenems was reported at a rate of up to 50% for patients with positive history and a positive skin test result to penicillins. The similarity in the chemical structure of these two groups likely explains this high rate of cross-reactivity. In spite of this, a more recent study

reported the rate of cross-reactivity between penicillins allergic cases and carbapenams as 9.2-11% (Mirakian et al., 2015).

Cross-reactivity with cephalosporins

The degradation processes of penicillins and cephalosporins are very different in the body. Therefore, cross-reactivity between penicillins and cephalosporins is less likely. The rate of cross-reactivity between penicillins and first and second generation cephalosporins is reported to be 8% to 10% (Mirakian et al., 2015).

1.9 The role of genes

There is growing evidence that genetic factors may play a significant role in beta-lactam allergy; however, there is a lack of knowledge regarding the genetic and epigenetic underpinnings of beta-lactam allergy as a whole. Evidently, having a better understanding of the genetic factors that are involved in beta-lactam allergy will improve our understanding of the mechanisms of the reaction. Therefore, having a clear view of the genetic factors involved can lead us to develop better preventive methods and strategies as well as effecting better drug design and treatment strategies in the future.

Up to the present, no consistently replicated reports of genetic variants associated with beta-lactam allergy have appeared. So far, 19 studies have been published on genetic predictors associated with beta-lactam allergies. There is a widespread recognition of the lack of knowledge in this field, despite its presumed effect on the individual's life and the health care system as a whole.

Among all the published studies there has only been one study which used a genome-wide approach, by conducting an association study using the Immunochip fine mapping array (Gueant et al., 2014). However it must be highlighted that this study was not a fully comprehensive genome-wide association study, because the Immunochip only comprehensively covers immune-related candidate genes. As is shown in Table 1.6, 12 studies suggested that pro-inflammatory cytokine genes such as IL4R, IL4, IL13 are involved in IgE mediated beta-lactam reactions (Oussalah et al., 2016).

Table 1.6 Published genetic association studies

First author, year, journal	Geographical region	Study design and approach	Number of Cases	Case defin-ition	Number of Con-trols	Control definition
(Huang et al., 2009, Eur J Clin Pharmacol)	China	Case–control (candidate gene)	242	BL allergy	240	Healthy controls
(Gao et al., 2008, Eur J Clin Pharmacol)	China	Case–control (candidate gene)	144	BL allergy	88	Healthy controls
(Gueant-Rodriguez et al., 2008, Pharmacogenomics J)	Italy	Case–control (candidate gene)	167	BL allergy	260	Healthy controls
(Apter et al., 2008, J Allergy Clin Immunol)	USA	Case–control (candidate gene)	23	BL allergy	39	Healthy controls
(Qiao et al., 2007, Eur J Clin Pharmacol)	China	Case–control (candidate gene	102	BL allergy	86	Healthy controls
Guglielmi et al., 2006, Allergy)	France	Case–control (candidate gene	44	BL allergy	44	Healthy controls

First author, year, journal	Geographical region	Study design and approach	Number of Cases	Case definition	Number of Controls	Control definition
[Yang et al., 2006, Chin Med J (Engl)]	China	Case–control (candidate gene	113	BL allergy	87	Healthy controls
(Gueant-Rodriguez et al., 2006, Pharmacogenet Genomics)	Italy	Case–control (candidate gene	210	BL allergy	265	Healthy controls
(Yang et al., 2005, Eur J Clin Pharmacol)	China	Case–control (candidate gene	158	BL allergy	89	Healthy controls
(Qiao et al., 2005, Allergy)	China	Case–control (candidate gene	245	BL allergy	101	Healthy controls
(Qiao et al., 2004, Allergy)	Korea	Case–control (candidate gene	448	BL allergy	101	Healthy controls
(Jean-Louis Gueant et al, 2015, Allergy Clin Immunology)	Spain	Case–control (Immuno-Chip	436	BL allergy	1218	Healthy controls

The table shows genetic association studies that reported genetic predictors in association with immediate-type hypersensitivity to BL antibiotics.

A case-control candidate gene study, based on self-reported beta-lactam allergy status, was conducted in 2007 with a title of "Clinical and genetic risk factors of self-reported penicillin allergy" (Apter et al., 2008). They identified 23 self-reported cases (predominantly white female) through the reports of 17 different allergist-immunologists. The allergic symptoms for the cases were reported as urticaria, angioedema or swelling, shortness of breath or chest tightness or wheeze after taking beta-lactam antibiotics. Out of 23 cases, 11 (48%) were

reported as having immediate reaction (within 2 hours of most recent intake dose), 11 (48%) were reported as having semi-immediate reaction (within 48 hours) and 1 (4%) was reported as having a delayed reaction. For 78% of the cases, the allergic reaction had occurred more than 5 years before enrolling in the study. They also recruited 39 healthy controls from the same cohort (Apter et al., 2008).

In this candidate gene study, IL4, IL4R and IL10 genes were selected for their previously reported role in IgE-mediated reactions. They also added LACTB gene for its known role in penicillin metabolism. They reported 3 significant SNPs in the L4 gene; rs11740584 (P=5.012), rs10062446 (P=5.021) and rs2070874 (P=5.035); and one SNP, rs2729835 (P=5.058), in the LACTB gene, as being marginally associated with penicillin allergy (Apter et al., 2008). The limitations of this study included small sample size, self-reported allergic cases and recruiting cases by different allergists (Apter et al., 2008).

Another case-control study was conducted involving 44 allergic patients (immediate reaction to beta-lactam) and 44 controls matched for age and sex (Guglielmi et al., 2006). This study showed no significant evidence for association between allergy patients and controls in the 15 SNPs studied (in STAT6, IL13, IL4RA, IL4, IFNGR1, IFNGR2 and FCERIB). However, sub-group analysis revealed significant associations in IL10 and IL4RA for immediate beta-lactam allergy in women with atopy (Guglielmi et al., 2006).

A further study based on a case-control sample of 210 cases and 265 age and gender-matched controls showed that genetic variants in the genes encoding

IL13 and IL4A were associated with the risk of immediate allergy to beta-lactam antibiotics (Gueant-Rodriguez et al., 2006).

A case-control study evaluating the association between tumour necrosis factor-α-308G>A with IgE-mediated allergy was conducted in 2006 (Gueant-Rodriguez et al., 2008). The study had a sample size of 427 subjects (167 cases, and 260 age and gender paired control). It showed an uncertain influence of this locus on the pro-inflammatory pathways on IgE-mediated hypersensitivity to beta-lactams, because the GG genotype was found to be associated with allergy risk whereas the AA genotype was found to be associated with higher levels of serum IgE (Gueant-Rodriguez et al., 2008).

A case-control study published in 2013 evaluated the association of polymorphisms in the NOD1 and NOD2 genes with beta-lactam allergy (Bursztejn et al., 2013). 368 Italian and 387 Spanish patients were compared with 368 and 326 controls, respectively. The Italian cohort revealed an association between a SNP in NOD2 and both higher risk of beta-lactam allergy and elevated levels of IgE, whereas the Spanish cohort revealed an association with beta-lactam allergy in a different NOD2 SNP.

Finally a recent study conducted a regional fine-mapping genome-wide association study of the genetic predictors of beta-lactam allergy (Gueant et al., 2014). They used the Immunochip custom array (which comprehensively covers immune-related candidate genes) for genotyping 436 allergic cases from Spain and 1218 paired control subjects. They performed the replication study in 299 allergic Italian patients and 362 paired controls. They found that immediate

beta-lactam allergy was influenced by genetic variants in HLA-DRA. Also they reported significant association with the SNPs of C5, ZNF300 and HLA-DRA/HLA-DRB5 genes.

The paper concluded that genetic variants of HLA-DRA and the HLA-DRA/HLA-DRB5 region were significant predictors of allergy to penicillins but not to cephalosporins. These data suggest genetic variation related to HLA type 2 antigen may play a crucial role in penicillin allergy.

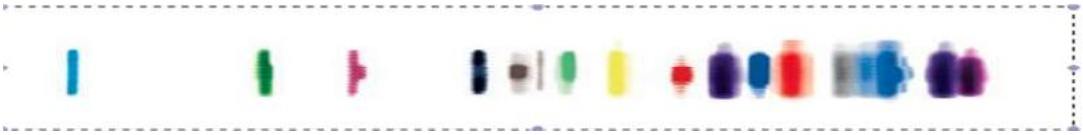
1.10 Motivation and aims of this thesis

There is a global importance of penicillin allergy to the health care system, related to the outcome of replacing beta-lactam antibiotics with second line drugs once the patient is labelled as having an 'allergy'. Firstly, there is the direct economic cost, since second line drugs are more expensive than first line beta-lactam antibiotics. In a health care system with finite resources this places great pressures on resources. Secondly, the patients' treatment plans may become limited as they may build up resistance to the second line drugs, which may have detrimental clinical implications as options of third line drugs may be limited and less effective. Thirdly, patients may be allergic to second line drug options, further limiting their management plans. Fourthly, the unnecessary overuse of second-line antibiotics will accelerate the evolution of bacteria resistant to these drugs, and will eventually render them useless.

Taking into account the complex aetiology of the beta-lactam allergy, the studies described in this dissertation aim to make genuine contributions to the field of pharmacogenomics by discovering biomarkers related to beta-lactam, allergy. By benefiting from having access to the large database of TwinsUK, and using the advantage of twin studies, I estimated the heritability of beta-lactam allergy and I also conducted a detailed genome wide association study and a metabolome wide association study among self-reported twins. There is a paucity of research in this highly significant yet complex field, which would therefore benefit from further data to help us to unravel the complexities in the field of beta-lactam allergy using genetic studies. This is important because of the potential impact on the clinical outcomes for patients, and also because there are wider economic implications. Genetic biomarkers provide a potentially beneficial route to better patient health outcomes, which makes this type of enquiry imperative.

The twin studies were followed up with a separate study on a newly recruited clinically-defined beta-lactam allergy cohort. Both studies used newly developed questionnaires to improve diagnostic accuracy.

Chapter 2: Beta-lactam allergy phenotypes and heritability in the TwinsUK cohort



2.1 Introduction

In the previous chapter, I reviewed the biology of allergy in a general context, covering drug reactions and diagnostic steps of allergic reactions to beta-lactam antibiotics. As I discussed in chapter one, previous studies suggest that both genetic and environmental factors may influence beta-lactam allergy. In order to understand and estimate the effect sizes of genetic variants in beta-lactam allergy, in this chapter we aimed to estimate beta-lactam heritability by using data from the TwinsUK cohort. Twin studies provide us with one of the most powerful designs for separating and quantifying the effects of genetic and environmental factors on common complex traits such as beta-lactam allergy.

2.1.1 The twin study design

Identical or monozygotic (MZ) twin pairs share 100% of their genome whereas non-identical or dizygotic (DZ) twins on average share 50% of their genome which is the same as normal siblings. Both MZ and DZ twins are assumed to share an equal environment; therefore a greater phenotypic similarity between MZ twin pairs compared with DZ twins is considered to be indicative of the role of genetic factors.

The classical twin design can be applied to study both quantitative traits, such as height, weight, blood count, and qualitative or binary traits such as disease status. For binary traits, when both members of a twin pair have the same trait, this is known as being concordant otherwise they are discordant for the trait.

When the pair-wise concordance rate among MZ twin pairs is significantly higher than that of DZ twins, it shows that there is a genetic contribution to a trait.

Currently there are no reported twin studies of beta-lactam allergy. However, previous twin studies on allergic diseases consistently show a significantly higher concordance rate in MZ pairs in comparison to DZ twin pairs. One of the largest and most comprehensive twin study on "allergic disease", with a sample size of 7000 twin pairs, was conducted by the Swedish twin registry. The study was based on self-reported allergy questionnaire data (Edfors-Lubs, 1971). Although the results of the study varied depending on the condition, overall the study demonstrated that environmental factors play a more important role in beta-lactam allergy than genetic factors.

Another twin study of atopy and allergic sensitisation markers was conducted on a sample size of 349 MZ and 533 DZ twin pairs and the result of this study confirmed the earlier findings on the importance of environmental factors (Strachan et al., 2001). Also they estimated a significantly higher concordance rate in MZ twins compared to DZ twins for hay fever, eczema and specific IgE positivity. Their estimates however did not apply to self-reported asthma (Strachan et al., 2001).

2.2 Materials and Methods

2.2.1 Study cohort (TwinsUK)

The sample set used in this heritability estimation study is known as the UK Adult Twin Registry or TwinsUK cohort. The UK Adult Twin Registry is a cohort started in 1993, by recruiting through media campaigns, in order to investigate conditions with higher prevalence in women such as osteoporosis and osteoarthritis. The success of the initial early studies, and the high success of using the twin model in the study of complex traits, led to the expansion of the cohort (Moayyeri et al., 2013). The UK Adult Twin Registry is run by the Department of Twin Research and Genetic Epidemiology (DTRGE) at St. Thomas' Hospital King's College London. Full ethical approval has been given for academic and commercial use (for more detail on ethical approval, please see section 2.2).

Twins are recruited without any selection for any particular disease or trait. More than 12,000 twins are registered on the TwinsUK research database and around 9000 of these twins are currently actively participating in research conducted by the Department of Twins Research and Genetic Epidemiology. The cohort contains approximately equal numbers of identical MZ (51%) and non-identical DZ (49%) twins aged 16 to 100. The cohort is predominantly female (80%), partly due to historical reasons and partly because female twins are more willing to volunteer than males.

All clinical, physiological and behavioural data are collected during visits to the Department of Twin Research and Genetic Epidemiology after providing patient information sheets (see Appendix A) and filling in consent forms (see Appendix B). Other health or lifestyle related self-reported information is collected by conducting questionnaires either during the visits or by post or email (Boomsma et al., 2002, Rijsdijk and Sham, 2002).

In order to determine the zygosity status of the participant twins in the TwinsUK adult registry, the 'peas in a pod' questionnaire (PPQ) (see Appendix C) is used for the majority of cases. Definitive zygosity scores in the peas in a pod questionnaire is an accurate assessment of the zygosity of the adult twins (Moayyeri et al., 2013). In the uncertainty after administering of the PPQ, additional tests are used to determine zygosity status in TwinsUK (Table 2.1). Several different genotyping and screening methods such as genome scan with more than 400 markers, DNA 5 markers, DNA 8 markers, DNA 10 Markers and DNA 16 markers are available (Moayyeri et al., 2013).

Table 2.1 Zygosity status determination on TwinsUK registry.

Inferred from co-twin/PPQ	Peas in a Pod Questionnaire (paper questionnaire, online, Phone or visit)
Self-report	Based on self-reported zygosity by twins during the registration with the DTRGE
DNA	DNA (8-10 markers)
Genome	Genome Scan (400+ marker test)
DNA 16 markers	DNA 16 markers done at KCL
DNA_47SNPs	OxfordDNA_47SNPs done at Oxford
DNA 5 markers	DNA 5 markers done at KCL

This table gives information on how the zygosity status of the participants is determined in TwinsUK adult registry.

Table 2.2 provides a summary of the TwinsUK adult registry and available samples. There are 10,393 FF (female/female) twins, 2,002 MM (male/male) twins and 328 FM (female/male) twins. In total there are 6,369 twin pairs. Over 1000 different clinical and biomarker data are available for most of the participants (Moayyeri et al., 2013).

DNA data for 7,548 twins and 995 parents are available, and 5,710 twins have genome wide genotyped data. DNA methylation data are available for 5,000 twins and next-generation sequencing data are available for more than 2000 twins (Moayyeri et al., 2013).

Table 2.2 The TwinsUK adult registry update. Based on information from (Moayyeri et al., 2013).

Name of register	**UK Adult Twin Registry (TwinsUK)**
Country	United Kingdom
Kind of ascertainment	Volunteers unselected
Opposite-sex twins (yes or no)	Yes
Number of pairs (Separated by birth range and sex)	1900-1920: 22 FF; 2 MM; 2 FM 1920-1930: 390 FF; 70 MM; 4 FM 1930-1940: 1,518 FF; 189 MM; 32 FM 1940-1950: 2,480 FF; 308 MM; 58 FM 1950-1960: 2,172 FF; 429 MM; 88 FM 1960-1970: 1,889 FF; 486 MM; 62 FM 1970-1980: 1,356 FF; 394 MM; 40 FM 1980-2000: 558 FF; 124 MM; 40 FM
Grand total	10,393 FF; 2,002 MM; 328 FM (6,369 pairs)
Major interests	Common complex diseases and ageing traits
Traits measured	Full questionnaires and clinical examinations on majority of twins for wide range of over 1,000 clinical and biochemical traits including: cardiovascular diseases, obesity, metabolic syndrome, respiratory diseases, dermatology, osteoarthritis, osteoporosis, eye diseases, coagulation system, immune function, cognitive function, gastro-intestinal system, pain thresholds, allergy, atopy, sexuality, pitch perception, and various aspects of personality.
DNA samples	13,458 aliquots from 7,548 twins 9,321 aliquots from 5,95 twins stored as back-up DNA samples from 995 parents and 1,227 siblings taken and stored

Other samples	119,511 blood samples (5,980 serum, 43,527 plasma EDTA, 10004 plasma Li heparin) from 7,681 twins (16,677 back-up samples) Range of 2-76 aliquots of various specimens at multiple time points
Comments	Monozygotic: Dizygotic ratio is approximately 1:1 Majority of twins are female with mean age of 55 years 5,710 twins with genome-wide association data 5,000 twins with DNA methylation data by the end of 2012 2,000 twins with next-generation sequencing data by the end of 2012 Data available for transcriptome across multiple tissues, telomere length, and metabolomics profile in different subsamples

2.2.2 Detail of ethics approval

Study title is "TwinsUK", the Research Ethics Committee reference is EC04/015, and the study approval date is Mar 2004. The body formerly known as the St. Thomas' Hospital Research Ethics Committee (REC) granted the ethical opinion for TwinsUK (see Appendix D). Following restructure and merging of REC, subsequent amendments were approved by the NRES Committee London – Westminster (see Appendix E).

2.2.3 Information on beta-lactam allergy from existing questionnaire data

The data used in this chapter were collected through a General Health Questionnaire (GHQ) administered to 3,755 monozygotic (MZ) and dizygotic (DZ) twins, aged between 16 to 82 years old, in the 6 year period between 2004 and 2010. The questions in the GHQ included two very simple yes/no questions related to allergy to penicillin. These questions were:

Are you allergic to penicillin? Yes/No

If so, have you had any treatment for your allergic reaction? Yes/No

These simple yes/no questions and did not provide information on the culprit drug (type of antibiotic that caused the allergic reaction), seriousness of the reaction or any available treatment/ medication, time delay of the reaction or type of the reaction.

Table 2.3 summarises the general characteristics of the twins participating in the General Health Questionnaire. The mean age of the participants was 53 years (52.9 ± 13.3) and ranged from 16 to 82. Data were obtained from 1694 DZ twins (individuals) and 2051 MZ twins (individuals) and 10 individuals with unknown zygosity (UZ). The male to female ratio was approximately 1:10, which is similar to the whole TwinsUK cohort.

Table 2.3 General characteristics of twins answering GHQ

	Visits from 2004-2010
Number of individuals	3755
Age	52.9 ± 13.3
Sex (male/female)	38 (1.0%)/3717 (99%)
Zygosity (DZ/MZ/UZ)	1694(45%)/2051(55%)/10(0.3%)

Age is presented as mean ± standard deviation.

2.2.3.1 Potential limitations of the original health questionnaire

There are a number of limitations of the General Health Questionnaire that may cause bias in our results:

- **Self reported data**

  The data from the GHQ is a self-reported data reported by twins. Therefore there is a lack of clinically confirmed allergy. We already know (see chapter 1) that up to 10% of people taking beta-lactam antibiotics report as being allergic to them but only 10% of these reported as being allergic after clinical examination.

- **Missing data**

  There are two main issues in terms of missing data in this study. The first problem occurs the yes/no question is unanswered. In this case we do not know whether a blank means "unknown" or whether it means "not allergic to beta-lactams". The second problem is reporting bias. If data are missing in a biased manner, for example if MZ pairs concordant for allergy to penicillin systematically failed to return the questionnaire – then this could have a serious effect on our results.

- **Lack of detailed information on beta-lactam allergic reaction**

  The original question in the GHQ is a simple yes/no question and did not provide any further information on the type of reaction, which is essential for our further studies. There is no information on clinically

approved beta-lactam allergy, type of reaction, time delay of the reaction or the culprit drug.

- **Recall bias**

  A "no" answer could be due to not recalling any mild allergic reaction, or in some cases even a more serious reaction.

- **Twins report pattern**

  Information in TwinsUK indicates that MZ twin pairs are more likely to cross-confer and give the same answer to the same question than DZ twin pairs. This will bias heritability estimates.

2.2.4 Estimation of the heritability from twin studies

We used Falconer's formula to calculate heritability (Hill and Mackay, 2004). Falconer's formula is used to estimate the genetic heritability of a trait based on the difference between twin correlations (Hill and Mackay, 2004).

$$\mathbf{h^2 = 2(r_{MZ} - r_{DZ})}$$

In this formula h^2 is the heritability, r_{MZ} is the monozygotic (identical) twin correlation and r_{DZ} is the dizygotic (non-identical) twin correlation in trait-sharing.

The confidence interval was calculated from the formulae below (Rijsdijk and Sham, 2002).

$SE(h^2) \sim 2*sqrt(SE(MZcorr)+SE(DZcorr))$

$SE(MZcorr) = (1 - MZcorr^2)/sqrt(n - 1)$

$SE(DZcorr) = (1 - DZcorr^2)/sqrt(n - 1)$

2.2.5 Design and administration of a new questionnaire for beta-lactam allergy

We sought to validate and expand the data used in TwinsUK general health questionnaire through the development of a more detailed questionnaire, specifically designed to extract clinically useful information from the twins who self-report suffering from allergic reactions to penicillins and/or other beta-lactam antibiotics.

The new questionnaire collected more specific information on the type of antibiotic that caused the allergic reaction, the symptoms experienced, family history of allergy, seriousness of the reaction, and the time interval between taking the medication and occurrence of the symptoms.

We sent the questionnaire to all 337 twin pairs (674 individuals) where at least one twin (co-twin) had answered 'yes' to the question of being allergic to penicillin in the GHQ, which included 19 male twin pairs and 318 female twin pairs with an age range of 19 to 78. Of these twin pairs, 30 MZ pairs and 18 DZ pairs were concordant for allergic status and 152 MZ pairs and 136 DZ pairs were discordant for self-reported allergy to penicillin.

2.2.5.1 European Network Drug Allergy (ENDA) questionnaire

As a basis for our new questionnaire, we first considered the European Network of Drug Allergy (ENDA) questionnaire. ENDA developed the first validated questionnaire designed specifically to help the diagnosis of drug hypersensitivity in 1999 (see Appendix F). Drug hypersensitivity reactions are difficult to diagnose due to a number of reasons including: large variation in the clinical picture of drug hypersensitivity reactions; lack of sufficient in-vitro and in-vivo tests; and weaknesses in the Gell and Coombs classification. A detailed history taking during clinical visits is the most important diagnostic step in order to determine the origin of the reaction. It is also an essential tool to verify the type and seriousness of the allergic reaction (for more details please see chapter 1) (Mirakian et al., 2015).

2.2.5.2 Development of the specific beta-lactam questionnaire

Although we would have preferred to use an already existing, clinically validated questionnaire such as the ENDA questionnaire, to validate basic information on

beta-lactam allergy status for the TwinsUK data this was not possible due to the difficulties in administering a complicated questionnaire full of clinical jargon to members of the general public. The ENDA questionnaire is long and contains difficult clinical terminology, as it was designed as a clinical tool. It is therefore not suitable as a questionnaire that can be completed by the twins without assistance. For these reasons we aimed to make our questionnaire as simple and as short as possible.

In developing the new questionnaire we took into account several factors: (1) the ENDA drug hypersensitivity questionnaire: (2) the British Society for Allergy and Clinical Immunology (BSACI) guidelines: and (3) a systematic review of the available literature. Most of the questionnaires we found were variations of the ENDA questionnaire and I BSACI guidelines. I also used the information and clinical history taking guidelines as described in Demoly et al. (1999, 2009 & 2010). From among those questionnaires, and under the supervision of Dr. M Rosario Caballero (Department of Respiratory Medicine and Allergy), I selected the most suitable questions for my study and its population.

We developed a single-sheet, 11-item questionnaire to collect clinically useful information on beta-lactam allergy status from the twins. The newly designed specific beta-lactam allergy questionnaire can be found in Figure 2.1. The questionnaire was designed to address the limitations of the penicillin allergy section on the general health questionnaire for TwinsUK and thus obtain better heritability estimation on beta-lactam allergy status in this cohort.

SUPPLEMENTARY DOCUMENT 2

Date:

KING'S College LONDON
University of London

ANTIBIOTICS ALLERGY QUESTIONNAIRE

1. Have you ever had any allergic reactions to any of following drugs? If yes, please mark which one

a. Penicillins: (Penicillin, Amoxicillin, Co-amoxiclav, Ampicillin, Flucloxacillin) **YES** ◯ **NO** ◯

b. Cephalosporins: (Cefalexin, Cefuroxime, Cefadroxil, Cefaclor, Cefuroxime, Cefotaxime) **YES** ◯ **NO** ◯

c. Other Drugs.....

If you have ticked **YES** ➔ please go to **Question 2**

If you have ticked **NO** you have completed the Questionnaire please return in the **FREEPOST envelope** provided. Many thanks for your cooperation.

PLEASE ANSWER ALL QUESTIONS TO THE BEST OF YOUR ABILITY.

2. For what infection was the antibiotic prescribed for?

Cold/ Flu ◯ Chest infection ◯ Dental procedure ◯ Surgery ◯

Other.....

3. Was the allergic reaction clinically diagnosed by a doctor (GP / Hospital DR)? YES ◯ NO ◯

4. Which antibiotic did you have the allergic reaction to?

Name of Drug..... Date of reaction.....

5. What were your symptoms? **(Multiple boxes can be ticked)**

Rash/Hives/Itchy	◯	Wheezing	◯	Throat tightness	◯
Swelling	◯	Coughing	◯	Dizziness	◯
Loss of consciousness	◯	Vomiting	◯	Diarrhoea	◯

Anaphylaxis (immediate, severe, potentially life-threatening allergic reaction) ◯

KING'S College LONDON
University of London

Other symptoms.....

6. How long after taking the antibiotic did you notice these symptoms?

Less than one hour ◯ After 1hour ◯ After 1 day ◯

7. Did you require any treatment to control the reaction? **(Multiple boxes can be ticked)**

No therapy ◯ Just stop taking the drug ◯ Antihistamines ◯ Corticosteroids ◯

Adrenaline ◯ A&E treatment ◯

8. How long after stopping the antibiotic or treatment, did the symptoms disappear?

Less than a day ◯ A week ◯ Two weeks ◯ More than 2 weeks ◯

9. Did the reaction occur the first time that you took the antibiotic?

Name of medication..... YES ◯ NO ◯

10. Have you taken any antibiotics since you had the ◯action? ◯ YES NO

If yes, which ones? .....

11. Does any member of your family have history of allergy to the same or any other kind of drugs/ or any other allergy? If yes please write further details below.

Relationship	Allergy
.....	.....
.....	.....

Figure 2.1 The newly designed specific beta-lactam allergy questionnaire for TwinsUK.
The questionnaire was designed to address the lack of information in the General Health Questionnaire on self-reported allergic reactions to beta-lactam antibiotics.

2.2.6 Using the new questionnaire to define improved definitions of beta-lactam allergy response

2.2.6.1 Study population for the new questionnaire

The new questionnaire was posted to a total of 337 twin pairs (674 individuals) from TwinsUK, where at least one of them had been recorded as being allergic to beta-lactam antibiotics according to the General Health Questionnaire. Our pool included 184 MZ (monozygotic) twin pairs and 152 DZ (dyzygotic) twin pairs and 2 UZ (unknown zygosity) twin pairs. In addition to the questionnaire, each participant received an invitation letter, explaining the purpose of the study (see Appendix G), and a pre-paid addressed envelope in which to return the questionnaire.

2.2.6.2 Collecting information from the new specific beta-lactam allergy questionnaire

Each questionnaire was reviewed using the history taking method (explained in detail in chapter 1) under the supervision of Dr. M. R. Caballero. It is important to bear in mind that, while each question was chosen in order to give us useful information in regards to the patient's allergy history, none of these questions on its own can give us a clear view. Rather, a combination of the answers to all questions gives us a much more clear picture of the allergy history, to allow us to review and decide on allergic status. We used this

approach to define: (1) a modified penicillin allergy status (yes/no); (2) a penicillin allergy seriousness status (mild/medium/severe); (3) a penicillin allergy time delay status (immediate/semi-delayed/delayed).

2.2.7 Modified penicillin allergy phenotypes – based on information from the new questionnaire.

1. Recalculating the allergy status heritability estimate, using information from the new questionnaire

After validating the TwinsUK beta-lactam data and defining the type and seriousness of the reaction, we recalculated the heritability in our cohort based on the new information. The new modified database contained data on 676 individuals who had been sent the allergy-specific questionnaire, plus the information for all the twins who just completed the General Health Questionnaire.

The heritability for the modified penicillin allergy status (yes/no) was calculated for twins who completed the new questionnaire and provided unambiguous answers. The results of our careful evaluation based on the new questionnaire were transferred over to the database for the full twins cohort, and our new data replaced the old data where it was different.

We re-calculated the heritability of the beta-lactam allergy in our cohort using Falconer's formula.

2. Heritability estimate for allergy seriousness

We also calculated the heritability for penicillin allergy seriousness (mild/medium/severe). These severity data were also imported to the dataset for the full twins cohort.

2.3 Results

2.3.1 Heritability estimation for penicillin allergy (yes/no) using the data from the General Health Questionnaire

Table 2.4 provides data for the heritability calculation based on the General Health Questionnaire, which was completed by 2115 twin pairs. The table indicates the number of 2-twin families which were concordant-YES, discordant, and concordant-NO, split by DZ/MZ zygosity.

Table 2.4 Number of concordant and discordant twin pairs based on the General Health Questionnaire.

Table of families with N=2 members				
	0=Concordant NO Twin pairs	**1=Discordant Twin pairs**	**2=Concordant YES Twin pairs**	**Total**
DZ	725 (59%)	148 (12%)	19 (1.5%)	892
MZ	1005 (82%)	183 (15%)	35 (3%)	1223

This table provides information on the concordant and discordant twin pairs for penicillin allergy status. MZ=monozygotic twin pairs; DZ=dizygotic twin pairs.

Table 2.5 shows the heritability estimation for penicillin allergy using the self-reported data from the General Health Questionnaire. The point estimate for heritability is 16%, but we note that the 95%CI for heritability includes 0, so there is no significant evidence (using a 5% significance level) for heritability in this trait.

Table 2.5 Correlation of DZ and MZ twins based on the GHQ.

	Corr or h2	N	SE	Lower 95%CI	Upper 95%CI
DZ	0.112	892	0.033	0.047	0.177
MZ	0.193	1223	0.027	0.139	0.247
H2	0.1626		0.086	-0.006	0.331

Note that these calculations display:

(1) Lack of power (most twin pairs are 0-0)

(2) A large number of discordant pairs, relative to the concordant-YES pairs. We only had 19 DZ twin pairs who were concordant YES and 35 MZ twin pairs who were concordant YES, compared to 148 and 183 discordant DZ / MZ twin pairs respectively (see Table 2.4).

2.3.2 Response to the first question in the new questionnaire

The results from question 1 ("Have you ever had any allergic reactions to any of the following drugs?" (penicillins/cephalosporins/other drugs)) are summarized in the tables below. This question helped us to get an idea if they had enough knowledge of the name or subfamily of the drug that they (or their co-twin) claimed to be allergic to. In terms of penicillin allergy, out of 676 twins

(individuals) who were sent the new questionnaire, 301 twins (individuals) responded as being allergic, with 220 confirming no allergy and 154 not answering the question. With cephalosporin allergy, only 16 twins (individuals) responded as being allergic, whereas 321 (individuals) responded as being not allergic. 334 twins did not respond to the question with 4 twins unable to remember (Table 2.6).

Table 2.6 Results for question 1 in the new questionnaire (penicillin/cephalosporin allergy)

	Yes	No	Don't remember	Not answered	Blank
Penicillin allergy	301 (44.5%)	220 (32.5%)	–	2 (0.3%)	152 (22%)
Cephalosporin allergy	16 (2.3%)	321 (47.4%)	4 (0.6%)	171 (0.25%)	163 (24%)

2.3.3 Results for Question 2 of the new questionnaire

The table below (Table 2.7) shows the results from question 2 in the new questionnaire ("for what infection was the antibiotic prescribed for?"). The table indicates that the majority of twins (38.1%) had antibiotics for reasons other than chest infection, cold or dental surgery. 27.2% of twins were prescribed antibiotics for an episode of chest infection, with only 10.3% taking antibiotics post dental procedure. The results also show 11.1% had antibiotics pre/post

surgery and only 7.7% were given antibiotics for cold and flu. 5.2% of participants did not respond to the question.

Table 2.7 Summary of responses to Q2 of the new questionnaire

Reason of antibiotic prescription	Response rate
Cold/Flu	7.7%
Chest infection	27.2%
Dental procedure	10.3%
Surgery	11.1%
Other	38.1%
Not answered	5.2%

2.3.4 Results for Question 3 of the new questionnaire

Out of 301 positive responses to question 1, 250 individuals reported they were diagnosed by a doctor at the time of the reaction. Only 37 twins reported not being seen by a doctor, and 10 twins did not respond to the question. Table 2.8

summarises the responses to question 3 ("Was the allergic condition clinically diagnosed by a doctor?").

Table 2.8 Summary of responses to question 3 of the new questionnaire

	Yes	No	Don't remember	Not answered
Allergic reaction diagnosed by doctor	250 (83%)	37 (12%)	2 (0.7%)	12 (3.9%)

2.3.5 Comparison of results between the old and new questionnaires and new questionnaire

Table 2.9 shows the results for a comparison of the data between the old (GHQ) and new questionnaires. The table shows that most not-allergic responses stayed as not-allergic (only 6 converted to being allergic). However, a large proportion of positive allergic responses converted back to being not-allergic (58 / 483 converted).

Table 2.9 Comparison of the results for the old (GHQ) questionnaire and newly designed specific beta-lactam allergy questionnaire.

	New questionnaire		
Old questionnaire	**Not allergic**	**Allergic**	**Grand Total**
Not allergic	4152	6	4158
Allergic	58	425	483
Grand Total	4210	431	4641

2.3.6 Heritability estimation for beta-lactam allergy using the data from the new questionnaire

Table 2.10 summarises the heritability estimation for beta-lactam allergy (either penicillin or cephalosporin) after administrating the more detailed questionnaire, including the 95% confidence interval.

The use of the new questionnaire led to an increased heritability estimate from 16% to 21%. Also the 95% confidence interval did not include 0, therefore indicating a significant result (at the 5% significance level).

Table 2.10 Heritability estimation for beta-lactam allergy from the new specific beta-lactam allergy questionnaire.

	Corr or h2	N	SE	lower95%CI	upper95%CI
DZ	0.112	892	0.0330	0.0477	0.177
MZ	0.220	1223	0.0272	0.1676	0.274
h2	0.2167		0.0856	0.0488	0.384

h2=heritability; DZ=dizygotic twins; MZ= monozygotic twins

2.3.7 Heritability estimate for allergy seriousness of the reaction

Finally, heritability was estimated for allergy seriousness (mild/medium/severe) in our data, assuming that the 3-level seriousness score could be treated as a quantitative trait. Table 2.11 indicates that the heritability estimate increased after adding more phenotypic detail to our estimation.

The 95% confidence interval was above zero, which indicated a significant result (at the 5% significance level).

Table 2.11 Heritability estimation for allergy seriousness.

	Corr or h2	N	SE	Lower 95%CI	Upper 95%CI
DZ	0.036	892	0.033	-0.029	0.102
MZ	0.157	1223	0.027	0.103	0.212
h2	0.241		0.0871	0.071	0.413

2.4 Discussion

In this chapter, I used the TwinsUK cohort to make additional observations regarding the epidemiology and heritability of beta-lactam allergy. There was limited phenotypic data (from the General Health Questionnaire) that related to beta-lactam allergy. We improved the information on beta-lactam allergy for the TwinsUK cohort through the design and administration of a new questionnaire, which was used to construct new variables relating to the seriousness and type (time delay) of the allergic response. I used this information to calculate heritability estimates for allergic phenotypes.

By using this new specific beta-lactam allergy questionnaire we were able to overcome some of the limitations of the old questionnaire. However, a number of limitations still remained:

- Self reported data subject to subjective bias.
- Because this was a postal questionnaire, co-twins were not blinded to the other twin's response.
- Some twins may not be allergic to beta-lactam antibiotics, whilst they may be allergic to another family of antibiotics.
- Most people do not have a clear understanding of drug allergy symptoms.
- Patients are not always able to recall history of prior drug exposure.
- Childhood history of allergy is common, but unfortunately patients can't recall any evidence of allergy.
- Sometimes it is unclear whether the symptoms are due to an allergic reaction or provoked by an underlying infection.

- In some cases, patients are taking more than one drug at the same time.
- Recall bias.

Using the available self-reported allergy data for 1223 twins from the GHQ, the heritability was estimated as 16% for this definition of beta-lactam allergy, which was not significant. However, following the administration of the new detailed questionnaire to the twins who reported being allergic plus their co-twins totalling 676 individuals, a re-calculation of the heritability estimate was performed. The result of this calculation showed a significant increase of heritability to 21%, which is positive evidence to show that our questionnaire was adding valuable information to the data.

Finally, to allow more specific heritability calculations, we considered seriousness and type of reaction for each participant and re-calculated the heritability for severity as a quantitative trait. The result shows an even more significant increase in the heritability estimation. This therefore also indicates that our careful examination of the questionnaire added more value to our data. Although the new questionnaire has not been fully validated this new questionnaire can be considered in the future as a better replacement for the current self-reported questionnaires. The detailed questionnaire provides us with better information, making it easier to identify false positive allergic reactions as rated by patients in the history-taking phase of the assessment.

Although there are some limitations to the new questionnaire (see methods section), the information on this questionnaire added more value to the TwinsUK beta-lactam allergy data. Another important contribution of this questionnaire

was that the heritability estimation increased after administering the more detailed questionnaire. This questionnaire also helped us to identify 58 non-allergic cases, who had self-reported positive allergy status in the General Health Questionnaire.

All our information for beta-lactam allergy in TwinsUK remains self-reported but we have reduced the number of the self-reported allergic cases which were likely false-positive, just by administering our new questionnaire. Many more questions could be added to this questionnaire in order to get a more accurate and detailed history of beta-lactam allergic reaction, but this would also have a downside, as it is likely that participants would be unwilling to complete a very long and complicated additional questionnaire.

Chapter 3: GWAS of beta-lactam allergy in the TwinsUK cohort

3.1 Introduction

A genome-wide association study (GWAS), also known as a whole genome association study, is a study that examines a large number of common genetic variations among different individuals to observe any variants associating with a trait (Stranger et al., 2011). This relatively new method searches the whole human genome for (typically) small variations known as single nucleotide polymorphisms (SNPs) (Bush and Moore, 2012). Each GWAS study is capable of looking for hundreds of thousands of SNPs at the same time to identify any genomic variation that may influence the risk of disease (Visscher et al., 2012, Wellcome Trust Case Control, 2007).

In case-control GWAS studies, participants are selected based on their phenotype for the specific trait or disease. Individuals with a particular disease are categorized as cases based on their clinical manifestations, and similar people without specified disease are classified as controls (Daly, 2010a).

As discussed, a GWAS is a hypothesis-free approach that offers the opportunity to overcome the complexity and difficulties that arise from a lack of information regarding the pathophysiology of the complex trait (Bush and Moore, 2012).

Power plays a significant role in a GWAS study (Bush and Moore, 2012). Although sample size is a major contributing factor to GWAS power, there have been successful GWAS studies on drug response even with small sample sizes (see chapter 1) (Cooper et al., 2008). For example, a GWAS study on warfarin yielded significant results on contribution of genetic variants in drug response a case/control sample size totalling only 181 (see chapter 1)(Cooper et al., 2008).

The success of such studies with small sample size in pharmacogenomics has encouraged the usage of GWAS in this dissertation.

Beta-lactam antibiotics are responsible for the most frequent cases of immediate reactions to drugs and the reactions can be life threatening. Most of these reactions appear within the first hour after drug intake (see chapter 1) (Mirakian et al., 2015). Beta-lactam allergy has a complex aetiology with multiple environmental and genetic causes (Apter et al., 2008). As explained in chapter 1, there have been candidate gene studies of beta-lactam allergy to identify causal variants and to prioritize genes and genomic regions. Yet, there is much that is still unknown about the aetiology of beta-lactam allergy (Gueant et al., 2014).

As discussed in chapter 2, estimates of heritability using beta-lactam antibiotic allergic cases showed that genetic factors were associated with this trait. Although the original beta-lactam allergy data for TwinsUK cohort was self-reported, administration of a new beta-lactam specific questionnaire to the same cohort helped us to exclude 58 cases from the original data, and the use of the new questionnaire led to an increased heritability estimate from 16% to 21% (see result section in chapter 2). By using the new questionnaire TwinsUK allergy data provided positive evidence of the involvement of a genetic component in beta-lactam allergy.

The factors contributing to running a GWAS on self-reported beta-lactam allergy using TwinsUK data include the proven involvement of genetic factors in beta-lactam allergy in the TwinsUK cohort via heritability estimation (refer to chapter 2). Other factors include having pre-existing genotyping data for the same

cohort, and previous examples of successful GWAS studies using self-reported data (e.g. Hinds et al 2013). Finally, there are previously reported examples of association between beta-lactam allergy and variants in genes such as interleukin-10 (IL10), IL13, IL4, HLA-DRA and TNFA (see chapter 1).

3.2 Materials and Methods

3.2.1 Study cohort

The sample set used in this GWAS study is known as the TwinsUK cohort. Pleas refer to Chapter 2 for further details of the TwinsUK cohort.

3.2.1.1 Details of Ethics approval

Study title: TwinsUK

Research Ethics Committee reference: EC04/015

Study approval date: March 2004

The committee formerly known as the St. Thomas' Hospital Research Ethics Committee (REC) granted the ethical approval for TwinsUK (see Appendix D).

Following restructuring and merging of separate RECs, subsequent amendments were approved by the NRES Committee London – Westminster (see Appendix E)

3.2.2 Subjects

A total number of 684 individuals were investigated in this study.

3.2.2.1 Case selection

After administering a detailed beta-lactam allergy specific questionnaire, 211 twins of European ancestry were identified as having self-reported beta-lactam allergy through the TwinsUK database at St Thomas' Hospital (see chapter 2). These individuals all had genome-wide genotype data available in the DTR database. Among the 211 cases, 181 (more than 85%) reported having their beta-lactam allergic reaction diagnosed by a doctor (Table 3.1). For this study, we made use of a specific consent form included with the beta-lactam allergy specific questionnaire administered (see chapter 2).

Table 3.1 Diagnostic status of self reported beta-lactam cases

	Allergic reaction cases
Diagnosed by doctor	**181(85.7%)**
Not diagnosed	**20(9.4%)**
Don't remember	**2(0.9%)**
Not answered the question	**8(3.7%)**

Based on question 3 of the specific beta-lactam questionnaire.

3.2.2.2 Control selection

We identified our control group from the same TwinsUK cohort. As a first step, we extracted those individuals with no history of beta-lactam allergy or any other drug reactions. To further refine the definition of controls, we then excluded individuals with any other allergy symptoms or reaction to any other substances. Thus our control group was a set of "super controls". In total, 473 individuals were identified as the control group for our TwinsUK GWAS.

3.2.3 Phenotype definition

We defined various phenotypes in our study cases;

1. Seriousness of the beta-lactam allergic reaction based on the information from the beta-lactam specific questionnaire (see chapter 2)

2. Time delay of response (type of reaction) based on the beta-lactam specific questionnaire (see chapter 2)

3.2.3.1 Seriousness of the reaction

From a statistical point of view, in cohort analyses, quantitative or semi-quantitative traits tend to be better-powered to detect a genetic effect. Taking this into account, we ordered our cases and controls based on the seriousness of

their type of reaction to beta-lactam antibiotics, to provide a supplementary analysis to the case-control analysis.

Cases were categorized based on the seriousness of the reaction by using the Brown classification (for more detail see chapter 1), into mild/medium/severe penicillin allergy responders coded as 1/2/3 respectively. For more information on the grading system please see chapter 2. Also a set of "non-allergic" individuals (controls) was recorded as code "0" (Table 3.2).

Table 3.2 Symptom seriousness status on allergic twins

Seriousness of the reaction	Number of twins
1	28(13.2%)
2	61(28.9%)
3	122(57.8%)
Total	211

1, Mild reaction; 2, Medium reaction; 3, Severe reaction
Based on question 5 in the specific questionnaire, followed by review by a Guy's allergy clinic consultant (see chapter2).

3.2.3.2 Time delay of the reaction

Cases were categorised into immediate/semi-delayed/delayed according due to the interval between drug intake and appearance of the first symptoms of beta-lactam allergic reaction. We used time-delay of the reaction which we also refer to as "type of the reaction", coded as 1/2/3 coded respectively (Table 3.3). Also the same set of "non-allergic" individuals (controls) were recorded as code "0" (See chapter 2 for more information on time delay).

Table 3.3 Type of reaction status on allergic twins

Type of reaction/time delays	Number of twins
1	14(6.63%)
2	43(20.37%)
3	154(72.9%)
Total	211

1, Immediate; 2, Semi-delayed; 3, Delayed reaction. Based on question 6 in the specific questionnaire followed by review by a Guy's allergy clinic consultant (see chapter 2).

3.2.4 Genotyping/ quality control/ imputation

3.2.4.1 Genotyping

The genotyping of the samples from the TwinsUK cohort was performed using a combination of Illumina arrays: HumanHap300, HumanHap610Q 1M-Duo and 1.2MDuo 1M (Richards et al., 2008, Soranzo et al., 2009). The normalised intensity data was pooled across four genotyping centres at the Centre National de Genotypage (France), Duke University NC (USA), Helsinki University (Finland), and the Welcome Trust Sanger Institute (UK). The genotype-calling algorithm Illuminus was used to assign genotypes. No calls were assigned if an individual's most likely genotype was called with less than a posterior probability threshold of 0.95. Validation of pooling was achieved via visual inspection of 100 random, shared SNPs for any overt batch effects.

3.2.4.2 Quality control (QC)

For quality control purposes, similar exclusion criteria were applied each of the four datasets (from the four genotyping centres) independently. Subjects were excluded based on genotype data when sample call rate <98%. SNP level QC measures included: minimum MAF (>1%MAF); minimum genotyping success rate (call rate >95%); test for HWE ($P > 10^{-6}$).

3.2.4.3 Imputation

Imputation was performed using the IMPUTEv2 software package (Howie et al., 2009). Two different reference panels PO (HapMap2, rel 22, combined

CEU+YRI+ASN panels) and P1 (610k+, including the combined HumanHap610k and 1M reduced to 610k SNP content) were used in data imputation.

Imputation quality score was thresholded at info>0.7, leaving a total of ~1.87 million SNPs post-imputation. In this analysis, only SNPs with minor allele frequency ≥0.02 were included in the analysis since the interest was to investigate common variants. (Small et al., 2011, Nag et al., 2014).

3.2.5 GWAS scan statistical analysis

As I explained in the previous chapter, when dealing with twin's data, MZ twin pairs are known to share nearly 100% of their genetic variation while DZ twin pairs are known to share 50% of their genetic variation, though the percentage may vary due to chance difference in sites of recombination among DZ pairs.

In order to process the TwinsUK GWAS data, we generated a genomic kinship matrix from the GWAS data, and used it to correct for the relatedness in the data. This matrix was used to correct for family structure by using linear mixed model packages. Linear mixed models (LMMs) such as GEMMA and GenABEL are widely used to test for association studies. Note that in addition to correcting for close relatedness between twins, this procedure will also correct for more distant relatedness (i.e. population stratification due to differences in ancestry among twin pairs).

There is an increasing interest in using LMMs due to their effectiveness in accounting for relatedness and incorporating population stratification when analysing twin's data (Eu-Ahsunthornwattana et al., 2014). The variety of different available LMMs raises the question of which package to use to analyse our GWAS data. A previous study compared different methods of LMMs for how they handled family-based data. This study suggested that there was no appreciable difference in power or in the control of Type I error between GEMMA and GenABEL, or indeed between any of the LMM packages they assessed. Overall, both the GEMMA and GenABEL packages are able to deal with either quantitative or binary data. We decided to use the GenABEL package for the analysing the genetically correlated twin data (Eu-Ahsunthornwattana et al., 2014).

GenABEL is an R library package for the analysis of binary and quantitative traits (http://www.genabel.org/). This package was developed to facilitate genome-wide association analysis of quantitative traits using data coming from extended families or collected from genetically isolated populations. Therefore this package is suitable for analysing the TwinsUK data. We treated our data as a linear quantitative trait so as to be able to run it with the GenABEL package.

3.2.5.1 GWAS scan tests

I applied four models to the TwinsUK data. Test 1 was a test of seriousness as a quantitative trait, and was applied to all 211 allergy cases plus 473 controls, in the hope of maximising power to find general signals. Test 2 was a test of

seriousness as a quantitative trait, applied to only the 57 immediate and semi-delayed allergy cases plus 473 controls, in the hope of maximising power to find signals specific to short time delay cases. Test 3 was a binary case/control test of allergy status, applied to all 211 allergy cases plus 473 controls, which removed the assumption of linearity implicit in the quantitative analysis of all cases (Test 1). Finally, Test 4 was a binary case/control test of allergy status, applied to only the 57 immediate and semi-delayed allergy cases plus 473 controls, which removed the assumption of linearity implicit in the quantitative analysis of short time-delay cases (Test 2).

3.2.6 Post-association analysis

Our post-association analyses on GWAS results involved the creation of QQ plots, Manhattan plots and regional association plots (LocusZoom plots).

QQ-plot

The quantile-quantile plot (Q-Q plot) is a graphical tool that compares the distribution of observed ordered negative log P-values against expected negative log P-values under the null hypothesis. This plot helps to indicate whether the study has achieved more significant results than expected under a global null hypothesis. The number of "independent" LD signals in the human genome is less than the number of SNPs assessed, and therefore the concentration band indicated in grey (indicating the 95% confidence interval under the global null

hypothesis) is wider than it should be (Wilk and Gnanadesikan, 1968, Ehret, 2010).

Manhattan plot

The Manhattan plot is a type of series plot used to present GWAS results. In a Manhattan plot, the x-axis shows the genomic position of SNP on chromosomes 1-22 and the y-axis shows the negative log p-value for association for each SNP (Ehret, 2010).

LocusZoom plot

The LocusZoom plot is a graphical tool to magnify the specific region of the Manhattan plot, which gives us detailed information on the genes and SNPs' locations in the region. The LocusZoom plot also displays regional information such as area and strength of the associated signal related to genomic position (Pruim et al., 2010).

3.3 Results

3.3.1 Results for Test 1: seriousness as a quantitative trait (all cases)

3.3.1.1 Q-Q plot

Figure 3.1 shows the Q-Q plot for the GWAS results on seriousness as a quantitative trait (Test 1). The result file contained 1453397 SNPs. Most of the SNPs adhere to the "null" (no significant difference between observed and expected p-value), The top 2 SNPs have p<5e-8, which is widely regarded as a true LD-corrected genome-wide significance threshold.

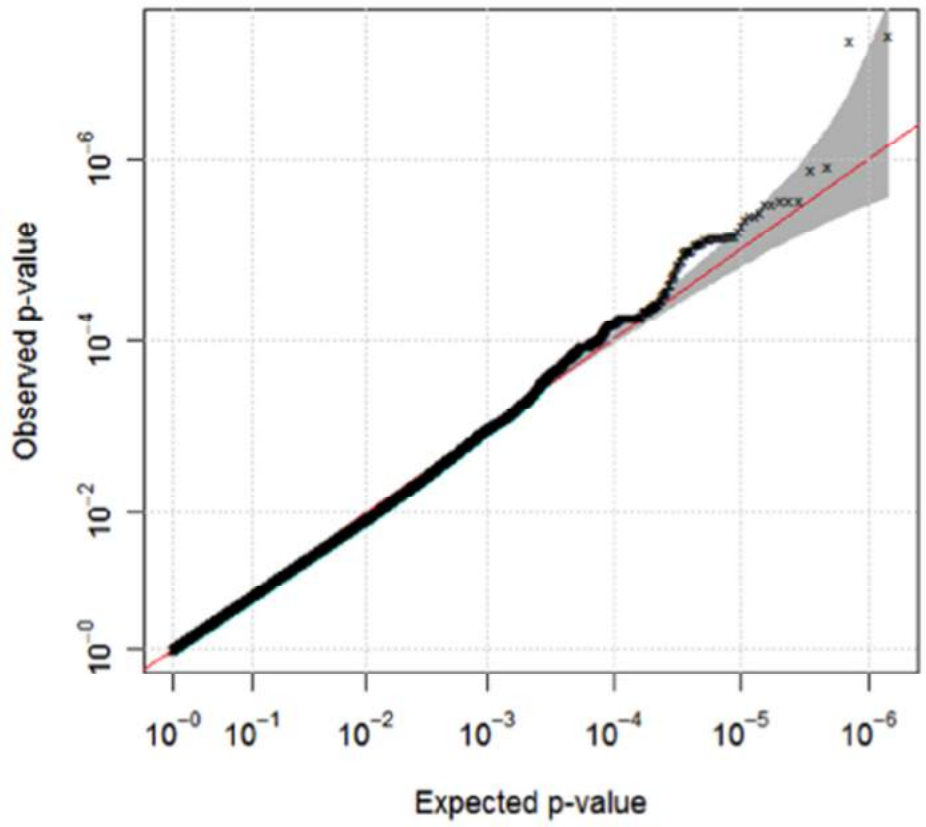


Figure 3.1 QQ plot on seriousness as a quantitative trait.
The negative logarithms of the observed (y axis) and the expected (x axis) ordered P-values are plotted for each SNP. The red line indicates the expected distribution under the null hypothesis. The grey shaded region represents the 95% concentration band. There is no genomic inflation in lower end of the plot. Hits are shown at the upper end of the plot.

3.3.1.2 Manhattan plot

Figure 3.2 shows the Manhattan plot of GWAS analysis on seriousness as a quantitative trait. The red line indicates the statistically acceptable threshold (p=5e-8). A significant hit with a p-value of p<5e-8 on chromosome 15 is seen.

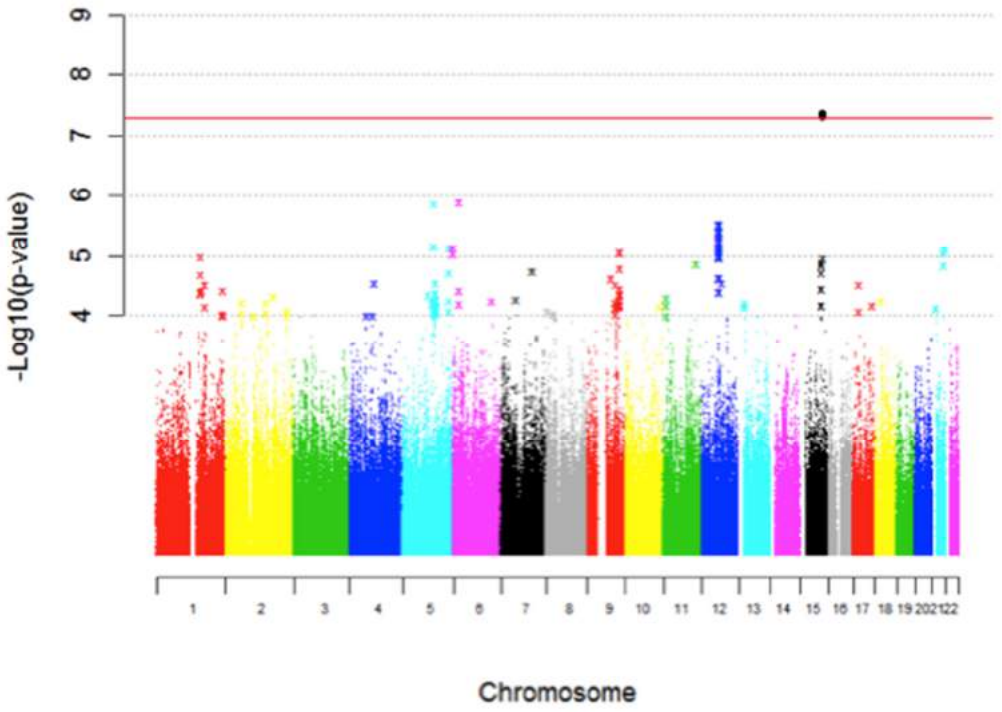


Figure 3.2 Manhattan plot for GWAS scan on allergy seriousness as a quantitative trait (all cases used).
The x-axis shows the chromosomal position and the y-axis shows the –log10p value. Chromosomes are shown in alternate colours. The red horizontal line indicates a threshold of genome-wide significance at a P-value of 5e-8. Each point represents a p-value for the SNP association test of the 211 case participants and 473 healthy control subjects.

3.3.1.3 LocusZoom Plot

Figure 3.3 shows a LocusZoom plot of the region containing the two SNPs with a p-value of p<5e-8. As shown in Figure 3.3, there is one typed and one imputed significant SNP supported by a number of SNPs underneath. MTHFS and BCL2A1 are two main genes under the signal, both of which have a plausible reason to have an impact on penicillin allergy (see Discussion).

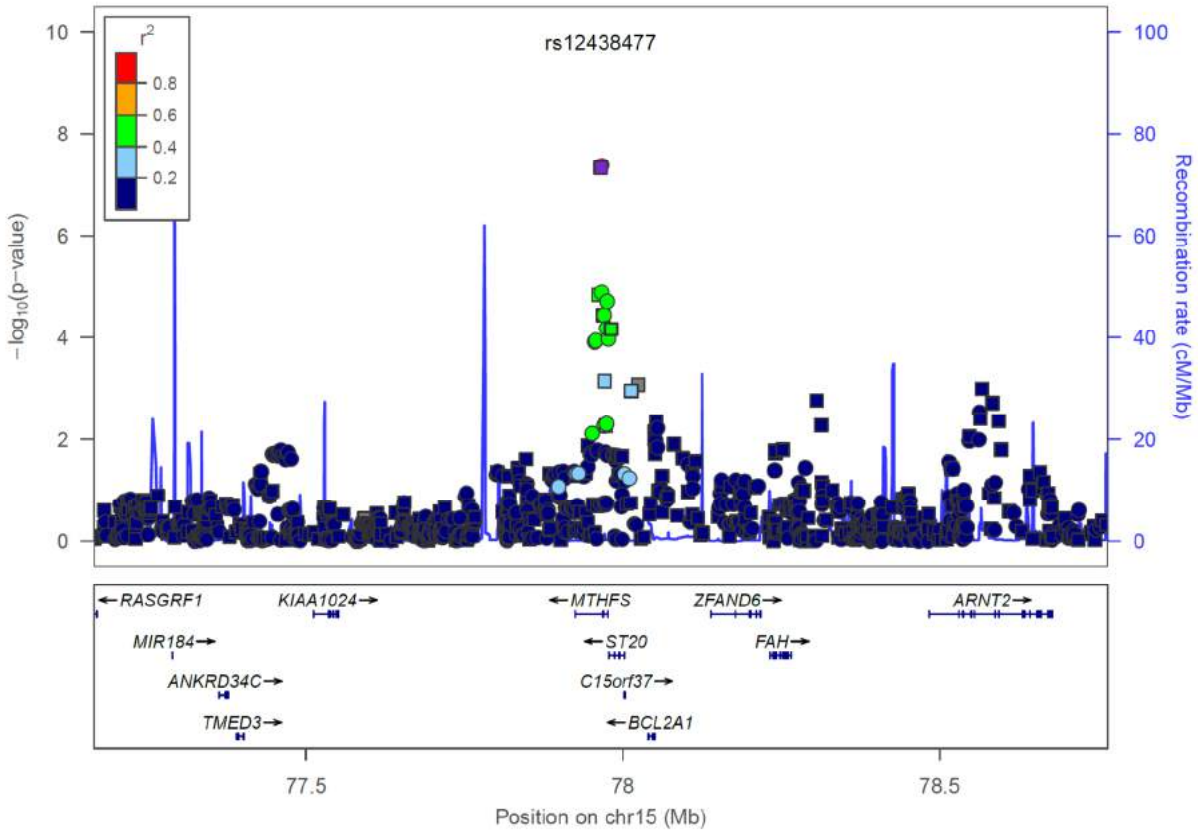


Figure 3.3 LocusZoom plot of region containing the significant hit for Test 1.
The main panel of a LocusZoom plot shows association p-values on the –log10 scale on the vertical axis, and the chromosomal position along the horizontal axis. Genes within the region are shown in the lower panel. Each filled circle represents the p-value for one SNP in the discovery cohort, with the top SNP rs12438477 shown in purple and SNPs in the region coloured depending on their degree of linkage disequilibrium (r^2) with rs12438477 (as estimated internally by LocusZoom on the basis of CEU (Utah residents of Northern and Western European ancestry) HapMap).

3.3.2 Results for Test 2: seriousness as a quantitative trait (short time delay only)

3.3.2.1 Q-Q plot

As shown in Figure 3.4, there are no significant results for allergy seriousness tested as a quantitative trait on short time delay (immediate+semi-delayed) cases (Test 2) and all the SNPs are in the grey area (concentration band indicating the likely region for "null" signals). No hits were seen with a p-value of less than 5e-8. The plot shows no evidence of population stratification, confounding effects, or systematic bias in the results from the statistical routines employed.

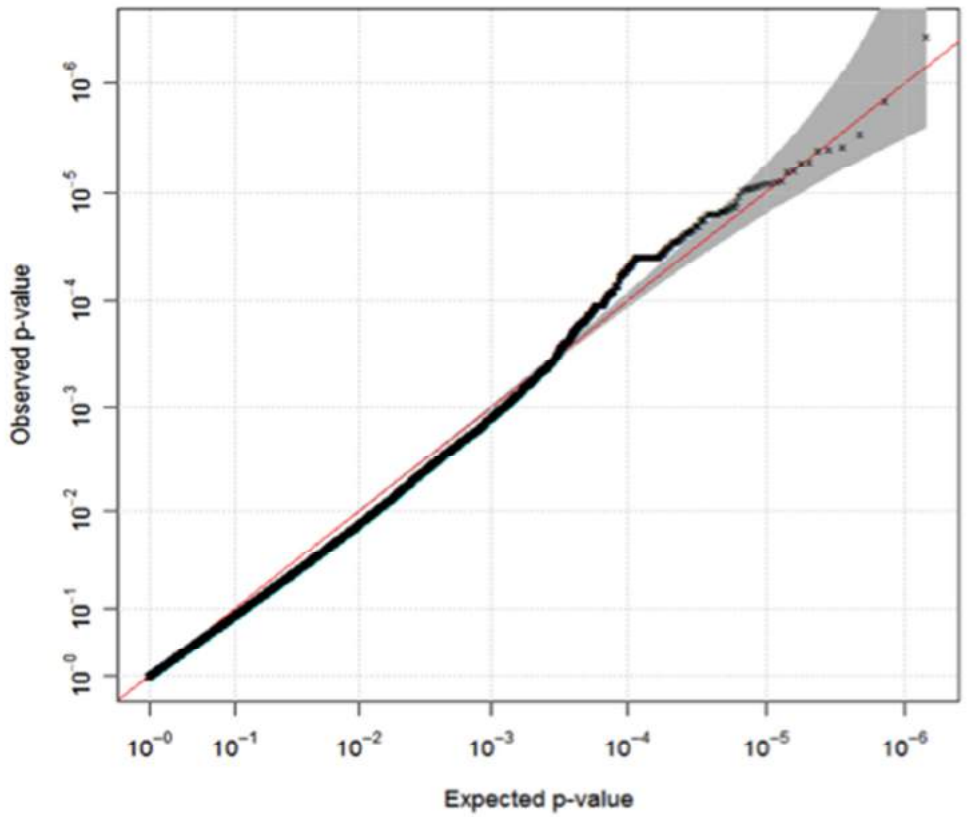


Figure 3.4 Q-Q plot of allergy seriousness as a quantitative trait (short time delay only).
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each SNP, and the red line indicates the expected distributions under the global null hypothesis. The grey shaded region represents the 95% concentration band. There is no genomic inflation in lower end, but nor are there any significant hits shown on the upper end of the plot.

3.3.2.2 Manhattan plot

As Figure 3.5 shows, the Manhattan plot of GWAS analysis on allergy seriousness status as a quantitative trait (immediate + semi-delayed cases only), confirms there are no significant hits with a p-value of p<5e-8.

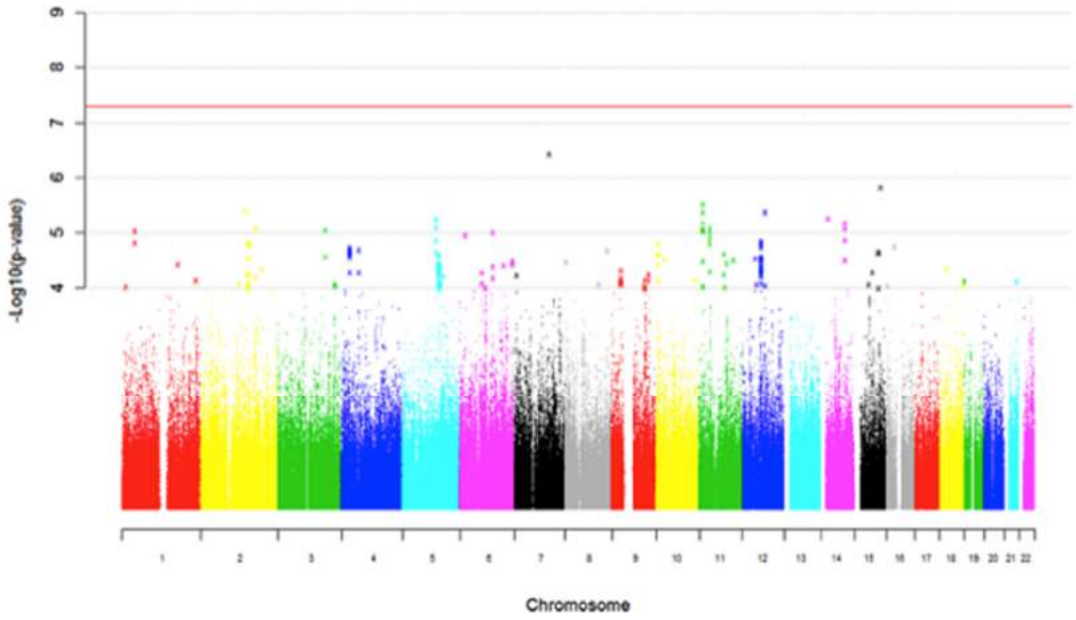


Figure 3.5 Manhattan plot for allergy seriousness as a quantitative trait (short time delay only).
The x-axis shows the chromosomal position and the y-axis shows the –log10 p-value. Chromosomes are shown in alternate colours. The red horizontal line indicates a threshold of genome-wide significance at a p-value of p<5e-8. Each point represents a p-value for the SNP association test of the 57 case participants and 473 healthy control subjects.

3.3.2.3 LocusZoom plot

As shown in Figure 3.6, there is still some evidence of a signal in the same region as the previously described hit (see result section for Test 1), but it is not a well-supported significant hit. The real SNP is still the same (rs12438477), but the p-value for the same SNP is reported as p>1e-5 in this analysis.

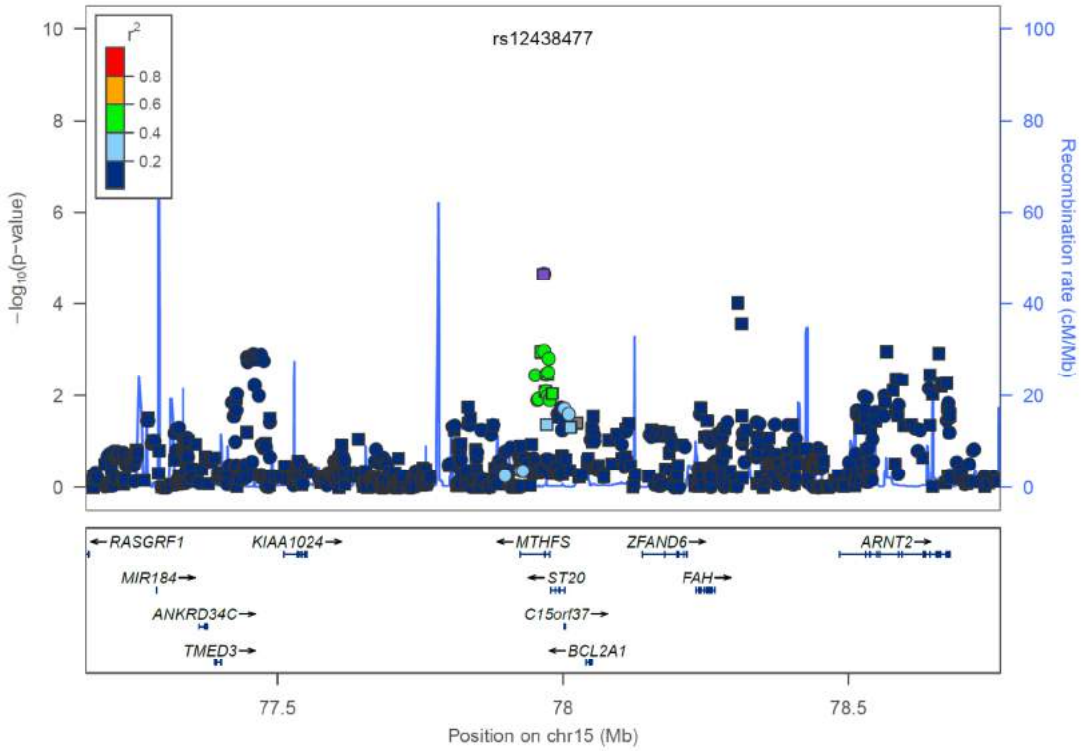


Figure 3.6 LocusZoom plot for allergy seriousness (short delay cases only). The main panel of a LocusZoom plot shows association p-values on the –log10 scale on the vertical axis, and the chromosomal position along the horizontal axis. Genes within the region are shown in the lower panel. Each filled circle represents the p-value for one SNP in the discovery cohort, with the top SNP rs12438477 shown in purple and SNPs in the region coloured depending on their degree of linkage disequilibrium (r2) with rs12438477 (as estimated internally by LocusZoom on the basis of CEU (Utah residents of Northern and Western European ancestry) HapMap).

3.3.3 Results for Test 3: allergy status as a binary case/control trait, all cases

3.3.3.1 Q-Q plot

As shown in Figure 3.7, there are no significant results for binary allergy status and all the SNPs are in the grey area (concentration band). No hits are seen with

a P-value of less than 5e-8, and there is no evidence for departure from a global null hypothesis.

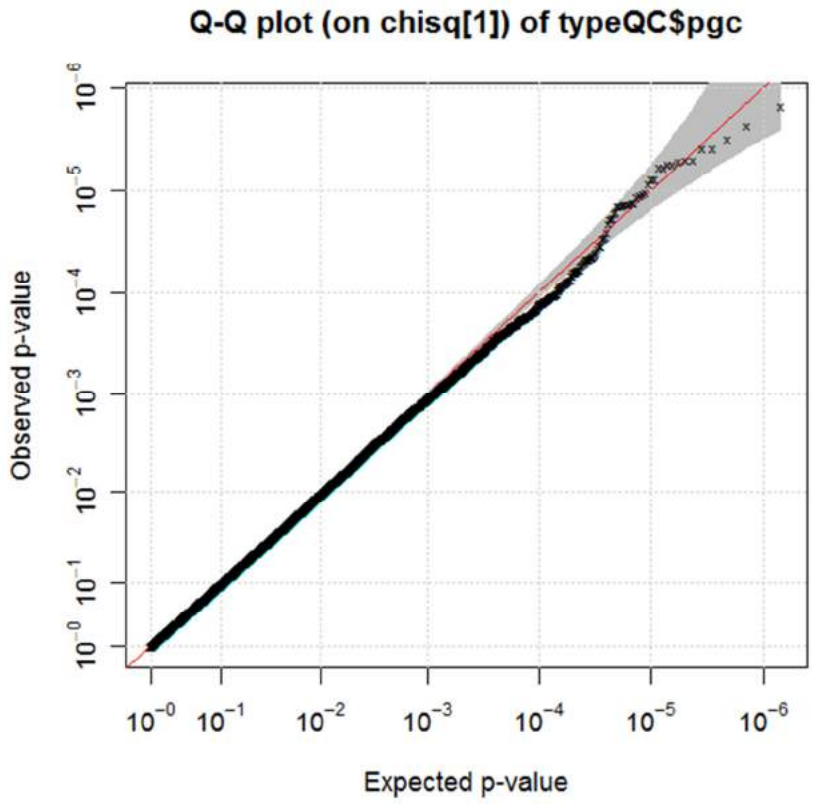


Figure 3.7 Q-Q plot for allergic response as a binary case/control trait. The negative logarithm of the observed (y axis) and the expected (x axis) p-value is plotted for each SNP, and the red line indicates the expected distributions under the null hypothesis. The grey shaded region represents the 95% concentration band. There is no genomic inflation in lower end, and there is no significant hit seen on the upper end of the plot.

3.3.3.2 Manhattan plot

As Figure 3.8 shows, the Manhattan plot of GWAS analysis on binary allergy status confirms that there are no significant hit with a p-value of p<5e-8.

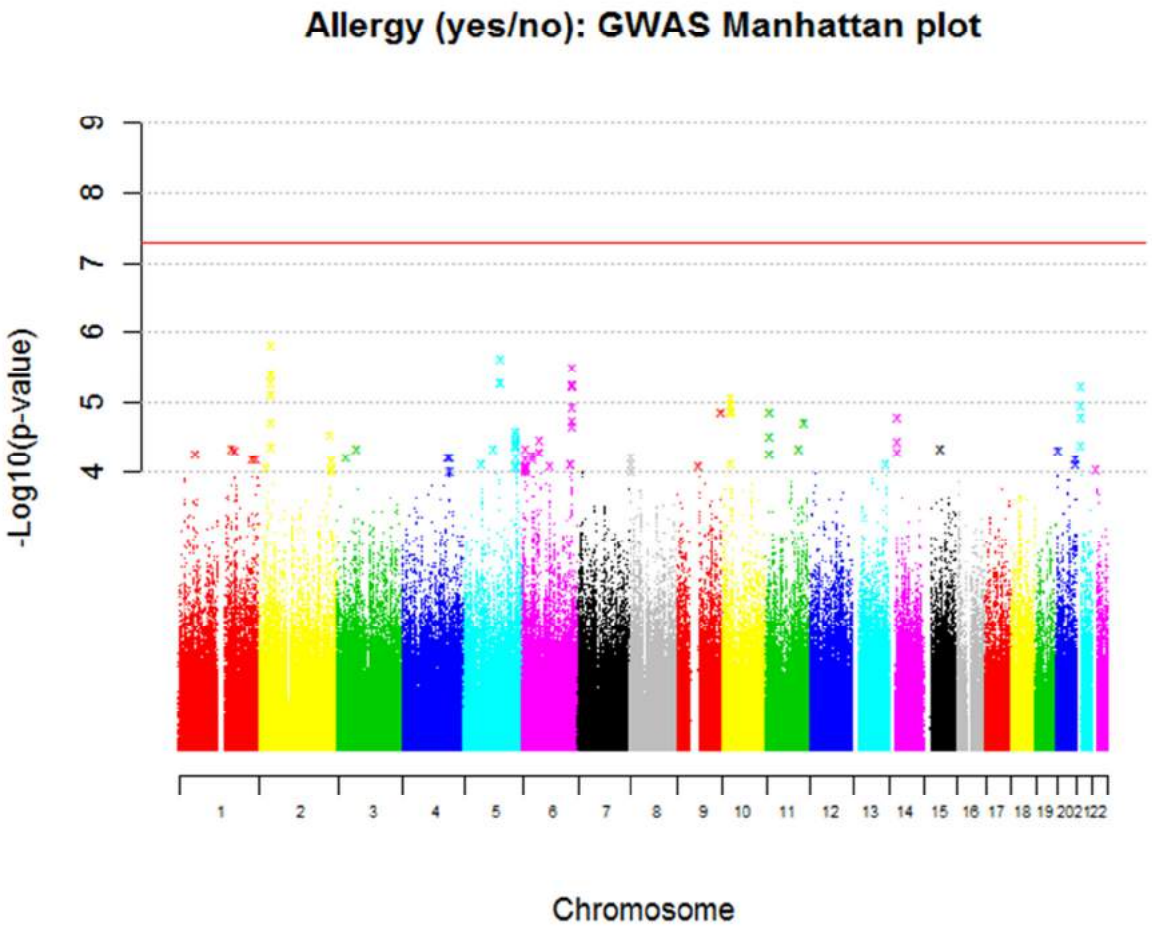


Figure 3.8 Manhattan plot of allergic response as a binary case/control trait.
The x-axis shows the chromosomal position and the y-axis shows the –log10p additive value. Chromosomes are shown in alternate colours. The red horizontal line indicates a threshold of genome-wide significance at a p-value of p<5e-8. Each dot represents a p-value for the SNP association test of the 211 case participants and 473 healthy control subjects.

3.3.4 Results for Test 4: allergy status as a binary case/control trait, short time delay cases only

3.3.4.1 Q-Q plot

As shown in Figure 3.9, there are no significant results for allergic response as a binary case/control trait (immediate + semi-delayed cases only) and all the SNPs are in the grey area (concentration band). No hits are seen with a p-value of less than 5e-8, and there is no evidence for departure from a global null hypothesis.

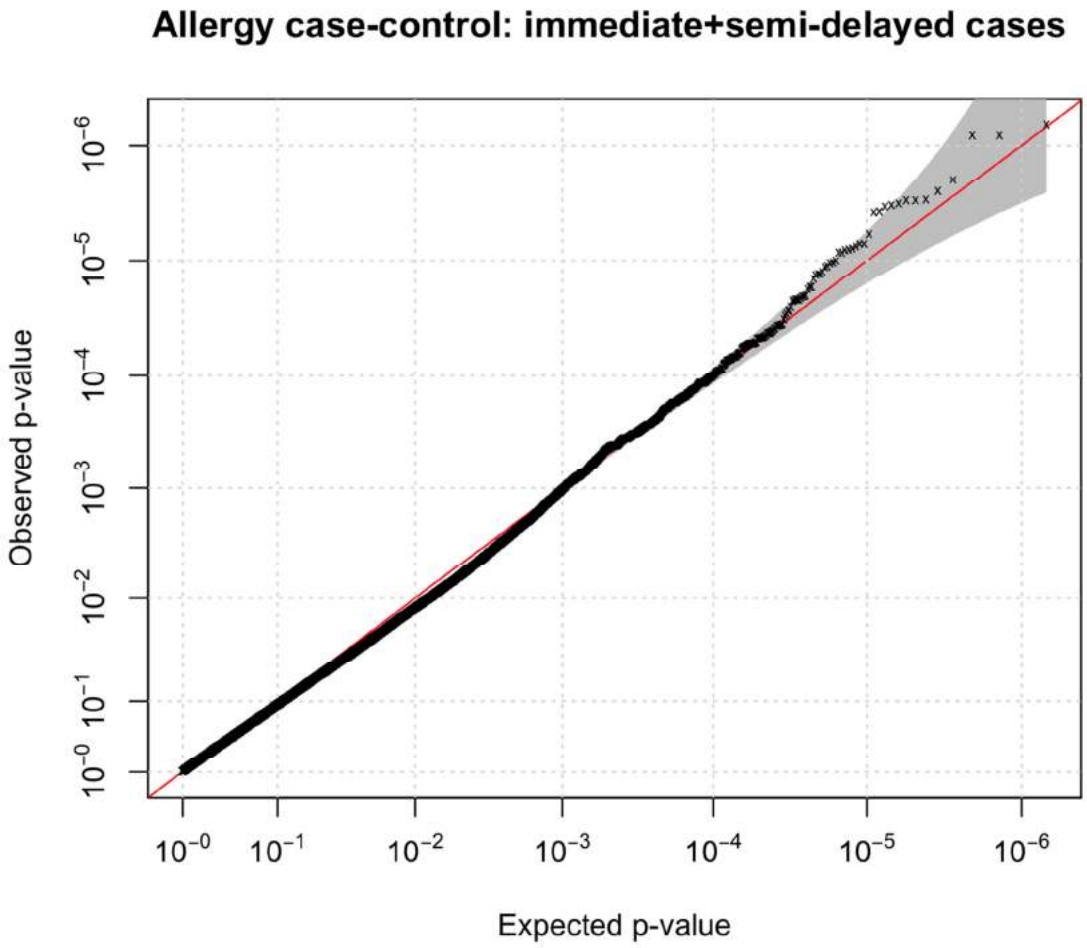


Figure 3.9 Q-Q plot for allergic response as a binary case/control trait (immediate + semi-delayed cases only).
The negative logarithm of the observed (y axis) and the expected (x axis) p-values are plotted for each SNP, and the red line indicates the expected distribution under the global null hypothesis. The grey shaded region represents the 95% concentration band. There is no genomic inflation in lower end, and there are no significant hits seen on the upper end of the plot.

3.3.4.2 Manhattan plot

As Figure 3.10 shows, the Manhattan plot of GWAS analysis on binary allergy status (short time-delay only) confirms that there are no significant hits with a p-value of p<5e-8.

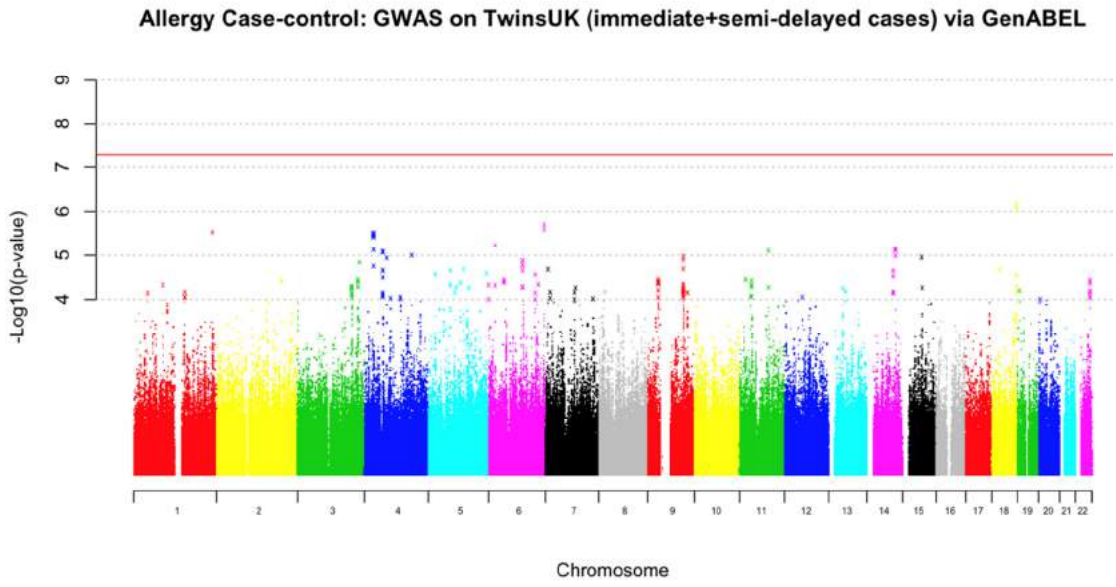


Figure 3.10 Manhattan plot of allergic response as a binary case/control trait on immediate and semi-delayed cases.
The x-axis shows the chromosomal position and the y-axis shows the –log10 p-value. Chromosomes are shown in alternate colours. The red horizontal line indicates a threshold of genome-wide significance at a p-value of p<5e-8. Each dot represents a p-value for the SNP association test of the 211 case participants and 473 healthy control subjects.

3.3.5 Checking the results of the previous studies in the clinical GWAS

We checked our result with the result of previously published studies. Table 3.4, shows that previously published results was not replicated in our TwinsUK study.

Table 3.4 Comparison of previously published finding within our data.

Gene	Variant	Chr	Pos	Rep in Twins GWAS	Notes
HLA-DRA	rs7192	6	32519624	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
C5	rs17612	9	1222765747	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
ZNF300	rs4958427	5	150258780	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs7754768	6	32528157	NO	Discovery hit in 2014 Immunochip study, but not genotyped in replication phase (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs9268832	6	32535767	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs2227139	6	32521437	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs2213586	6	32521072	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs2213585	6	32521128	NO	Candidate in candidate gene study (Gueant et al., 2015)

HLA-DRA	Rs7195	6	32520517	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA	rs8084	6	32519013	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
NOD2	rs2066845	-	-	NO	Candidate in candidate gene study
NOD1	rs2907749	-	-	NS	Candidate in candidate gene study (Oussalah et al., 2016)
IL4	Candidate gene	5	-	NO	Gueant-Rodriguez et al., 2006).
IL4R	Candidate gene	-	-	NO	(Gueant-Rodriguez et al., 2006).(Huang et al., 2009)
IL10	Candidate gene	-	-	NO	(Guglielmi et al., 2006).
IL13	Candidate gene	5	-	NO	(Gueant-Rodriguez et al., 2006).
LACTB	Candidate gene	-	-	NO	(Apter et al., 2008)
TNFA	Candidate gene	-	-	NO	(Cornejo-Garcia et al., 2012)
STAT6	Candidate gene	-	-	NO	(Huang et al., 2012)
IL18	Candidate gene	-	-	NO	(Ming et al., 2011)

Table showing the comparison of previously published findings within our data. All the significant SNPs and candidate genes in previous beta-lactam allergy studies have been checked for replication in all our scans (all-cases clinical-cases scan, amoxicillin-only clinical-cases scan, TwinsUK scan).

3.4 Discussion

As reviewed in Chapter 1, beta-lactam allergic response is an important health risk. Furthermore, incorrect labeling of patients as "allergic responders" contributes to greater use of second-line antibiotics, accelerating the evolution of bacterial resistance to these antibiotics. There is currently no comprehensive understanding of the relative importance of genetic and/or environmental factors that are associated with the elevated risk of allergic reactions to beta-lactam antibiotics (Mirakian et al., 2015).

Previously published case/control genetic association studies on beta-lactam allergy have been reviewed in Chapter 1. Nearly all case/control studies on beta-lactam allergies that were conducted on cases with a history of IgE mediated reactions (immediate reaction) were supported with clinical diagnoses (either Skin test, specific serum IgE or provocation test). However, patients with self-reported allergic reaction are still a big concern for our health care system due to the fact that they are not prescribed first-line antibiotics.

As discussed in chapter 2, heritability estimates provide positive evidence of a genetic component in beta-lactam allergy. Nevertheless, a phenotype derived from a self-assessed postal questionnaire cannot be considered as reliable as a clinical phenotype, and this remains the most important limitation to the TwinsUK GWAS. The fact that these heritability estimates improved (and became significant) when more carefully-defined phenotypes were used from the second

questionnaire (new specific beta-lactam allergy questionnaire) demonstrate the benefit of better phenotype definition.

I aimed to evaluate novel risk factors for beta-lactam allergy, especially short-term reactions. I used the self-reported beta-lactam allergy data from the adult TwinsUK cohort to carry out a genome-wide association study (GWAS) to identify common variants associated with variability to beta-lactam allergy among the allergic and non-allergic twins. The TwinsUK cohort consists of over 12000 volunteer twins. The cohort has been extensively molecularly characterized, including genome-wide genotyping and small molecule metabolic profiling via mass spectrometry. Using a custom designed questionnaire (see chapter 2), we defined more precise phenotypes; (1) seriousness of the beta-lactam allergy response and (2) type of reaction (time delay of response).

In the first GWAS scan (Test 1) we treated severity as a quantitative trait, and we used all cases regardless of time delays for their allergy status. This gave us an advantage of having more cases, which we hoped would increase the power of the GWAS. As previously outlined, all types of allergic reactions have a commonality in their biological pathways. Thus, combining all types of allergic reaction could also increase the power of the study, and any significant hits would in principle be those belonging to the common pathway.

This GWAS scan succeeded in providing a genetic hit in the MTHFS/BCL2A1 region that should be followed up in future studies, particularly in light of the role in immune function played by BCL2A1, which is a direct transcription target of NF-kappa-B in response to inflammatory mediators. MTHFS also has an

important role in inflammatory responses as it is involved in folate synthesis and is expressed in T-cells (Vajtr et al., 2014). Likewise, BCL2A1 has a significant function as it is involved in apoptosis and is directly regulated by NF-kappa-B in response to inflammatory mediators (Vogler, 2012).

The other GWAS scans (Tests 2-4) did not produce any significant results. In the case of Tests 2 and 4, this is likely due to the very small number of cases available when restricting to immediate+semi-delayed cases only (n=57). In the case of the binary case/control scan on all cases (Test 3), this is likely due to the reduction in power induced by removing the quantitative information on response seriousness from the test.

3.4.1 Strength and limitations of our GWAS scans

Here I discuss some of the strength and limitations of our GWAS scans, in light of the positive result that we obtained for Test 1.

Limitations include: (1) small case sample size (reduces power); (2) self-reported data via questionnaire (in chapter 2, I discussed limitations and bias related to data collection using the self-reported questionnaire); (3) recall bias as there were a large number of elderly participants in the cohort who may have found it difficult to recall the original reaction; (4) co-twin report bias (twins are more likely to report themselves as having a trait if one of the co-twins has been diagnosed with a trait); (5) **h**eterogeneity in the phenotype (combining cases with all types of reactions (all time delays), could possibly introduce

heterogeneity as a result of having different biological pathways amongst our cases); (6) misclassification of cases (it is possible that patients are misclassified with respect to penicillin allergy- it is known that the majority of patients who consider themselves to be allergic to beta-lactam are not); (7) lack of power calculation (we did not run a power calculation for TwinsUK GWAS, due to complications on power calculations in a mixed-model analysis and a lack of easy-to-use software).

Strengths include: (1) use of super-controls (not allergic to anything); (2) diagnostic status for many cases (although we identified our case group based on self-reported questionnaire, 181(85.7%) of them were reported as being diagnosed by a doctor); (3) case/control selection from the same cohort (reduces possibility for bias); (4) minimised the batch effect (DNA collection and genotyping were done in the same facilities).

As previously explained, many factors such as small sample size, batch effects and self-reported data are limitations of this study. Taking into account the small case sample size in our study and self-reported data, as a final step, we decided to check all the positive findings in the previously published studies and see if we could replicate those results, even if there was only a weaker (non-genome-wide-significant) signal in our study. However as shown in Table 3.4, previously published results was not replicated in our TwinsUK study. As previously discussed, most of the previous published GWAS studies have used a single locus analysis strategy. The locus containing our positive hit has not been assessed in previous candidate gene studies.

To conclude, the best result from the TwinsUK GWAS scan was the single hit at the MTHFS locus, which achieved genome-wide significance (p,5e-8). Knowing all the biases of self-reported data, and different biological pathways for immediate and delayed beta-lactam related allergic reaction, the question of whether this GWAS hit could be a false positive or not still remains. Below I list some supporting evidence to support the case that the signal from TwinsUK GWAS scan is not a false positive;

1. The local association plot looks good, with the lead SNP "supported" by several nearby SNPs in high LD (see Results section).
2. The p-value for the lead SNP is genome-wide significant, meaning that generating this signal by chance would be very unlikely unless there was some systematic bias at work here (see Results section).
3. The QQ plot shows no evidence of genomic inflation, suggesting there is no systematic bias in the data.
4. Systematic bias is unlikely. The genotype of this lead SNP would have been unknown to the person responding to the questionnaire. It is therefore hard to conceive a mechanism whereby the genotype influenced the person's response on the questionnaire, unless it was because that the genotype genuinely influences the outcome.

This was perhaps the most promising single result from this thesis, but until it can be replicated it must still be regarded as provisional.

Chapter 4: MWAS of beta-lactam allergy in the TwinsUK cohort

4.1 Introduction

Among a wide variety of antimicrobial drugs, beta-lactams are the most frequently prescribed antibiotics by clinicians, and the allergic reactions to these antibiotics are a major factor limiting their use (Mirakian et al., 2015). As discussed in chapter 1, genome-wide association studies have identified several risk loci associated to beta-lactam allergic reaction. However, many of these signals have yet to be confirmed in large, well-powered studies and the underlying biological processes are still largely unknown.

The recent ongoing revolution in "omics" technologies, such as transcriptomics and metabolomics, has enabled the deeper study of complex traits and is expected to provide more detail on affected pathways and more directly related information on the aetiology of the complex traits (Ma and Lu, 2011). The term "metabolomics" was first introduced in 2001 as a comprehensive and quantitative analysis of all metabolites in a biologic system (cell or body fluid) (Petersen et al., 2014).

4.1.1 Metabolites

Metabolites are known as products of enzyme-catalysed reactions, which take place naturally within all cells (Nicholson and Lindon, 2008). In other words metabolites are small-molecular-weight substances existing in all tissues and body fluids such as plasma, lymph, serum, urine, bile and sweat. Nuclear Magnetic Resonance Spectroscopy (NMR), Mass Spectrometry (MS) and Fourier

Transform Infra-Red Spectroscopy are the main metabolite measurement techniques used in metabolomic studies (Nicholson et al., 2002).

There are certain criteria which any compound meet before it can be classified as a metabolite (Nicholson and Wilson, 2003);

- Metabolites are compounds found inside cells
- Metabolites are recognized and acted upon by enzymes
- The product of a metabolite must be able to enter into a subsequent reaction
- Metabolites have a finite half-life, and they do not accumulate in cells
- Metabolites must serve some useful biological function in the cell

Investigation of the metabolic profile of humans provides useful information on both influential genetic factors and non-genetic factors such as dietary, lifestyle, age, nutrition habits, smoking and alcohol consumption, physical activity and environmental exposures (Nicholson and Wilson, 2003, Kastenmuller et al., 2015). The analysis of this information is expected to provide a new field of novel discoveries and understanding of all biological pathways (Nicholson and Wilson, 2003). Genetic factors usually cause a permanent change in protein function whereas environmental and physiological factors cause a more temporary impact on body metabolites (Ma and Lu, 2011). The effect of non-genetic factors on metabolite levels often reverts to a baseline state after the factors are removed (Kastenmuller et al., 2015).

At the molecular level, any genetic variations in the coding or regulatory regions of a gene can affect the structural or functional change of the target protein or its expression level; all of which can cause changes in metabolite levels. Furthermore, genetic polymorphisms in drug metabolizing enzymes and transporters can lead to changes in pharmacokinetics (Kastenmuller et al., 2015). As discussed previously, our understanding of the allergic reaction to beta-lactam antibiotics remains incomplete.

4.1.2 Beta-lactam metabolites and MWAS

As I described in chapter 1, hypersensitivity reactions to beta-lactams are due to different degradation products produced in vivo when the antibiotic is administered. Allergic reactions mainly involve ring-opened penicilloic acid molecules which bind covalently to proteins (Mirakian et al., 2015). To better understanding of beta-lactam metabolite formation in the human body, it is necessary to review the chemical structure of these antibacterial drug families once more.

As explained in more detail in chapter 1, all beta-lactam antibiotics contain a beta-lactam ring in their molecular structure (Pichler et al., 2011). Overall, the chemical structure of all beta-lactam families is similar and the distinction among them is based on the nature of the attached side-chain. After the administration of penicillin, the beta-lactam ring opens during nucleophilic attack, which results in the formation of penicilloyl metabolites (Solensky, 2003, Rosario and Grumach, 2006).

These metabolites are relatively small molecules and they cannot induce any immune response unless they are covalently bound to a protein (Mirakian et al., 2015). The covalent binding of these small metabolites to tissue macromolecules (proteins) generates the hapten-protein complex, which induces immune response. In IgE mediated beta-lactam allergic reactions, the presence of the hapten-protein complex is necessary for the formation of the complex antigenic determinant (Suhre et al., 2011).

In metabolomic genome-wide association studies (mGWAS), metabolite quantitative trait loci (mQTLs) are identified by associating SNPs with metabolite levels (Kastenmuller et al., 2015). In contrast, a metabolome-wide association study (MWAS) seeks to associate metabolite levels with a high-level trait of interest. For example, a recent MWAS of asthma showed that the metabolites from several pathways could distinguish asthmatic cases from controls, and also severe asthmatic cases from non-severe asthmatic cases (Singh et al., 2013).

In chapter 3, we described a significant hit associated with self-reported beta-lactam allergy in the Twins UK cohort. In view of the availability of metabolomic data for the same cohort, we decided to run a hypothesis-free non-targeted MWAS scan to identify new biomarkers or a metabolomic pathway associated with the GWAS signal.

Metabolites sampled from an individual years after a reported allergy attack are of interest for two reasons: (1) genetic differences affecting allergy risk are a life-long effect (genotypes never change), and may manifest themselves permanently

in different metabolite levels; (2) some allergic responses (delayed reactions) are long-lasting (e.g. changes to certain immune cell populations), and these may also manifest themselves in long-term differences in metabolite levels.

4.2 Materials and Methods

4.2.1 Case and control selection

In chapter 3, we identified a number of 211 self-reported beta-lactam allergy cases from the TwinsUK allergy-response data. Metabolomic data were in principle available for all our selected cases. A total number of 1023 non-allergic individuals with available metabolomic data were identified as controls. A total of 1234 individuals were investigated in this metabolome-wide association study (for more details please see chapter 3).

As previously mentioned, the case and control selection for this study was exactly the same as case/control selection for the GWAS study (chapter 3), but when we checked for data availability for the same case/control group, unfortunately MWAS data were not available for all individuals. This explains the difference in the number of cases and controls for the GWAS and MWAS studies.

4.2.2 Phenotype definition (seriousness and time delay)

Phenotype definitions are as described in Chapter 3 (please see section 3.2.3).

4.2.3 Metabolomic profiling and measurement

The Metabolon data in the TwinsUK registry consists of metabolomic data for 6055 twins, and these data are available for 510 different metabolites in total (Suhre et al., 2011). Metabolic profiling was done on fasting serum, by using non-targeted gas chromatography-mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS) (Suhre et al., 2011).

TwinsUK blood samples were collected after at least 6 hours of fasting. After keeping samples at 4°C for 40 minutes, samples were centrifuged for 10 minutes at 2,000g. After the centrifuging process, serum was removed from the centrifuged tubes. Four aliquots of 1.5 ml were placed into skirted micro-centrifuge tubes and stored at -45°C until sampling (Illig et al., 2010).

Metabolon data were collected over 173 days (35 measurements per day). As shown in Table 4.1, Metabolon data were divided into three batches; batch 1 (serum samples), batch 2 (plasma samples) and batch 3 (plasma samples) (Suhre et al., 2011). Batch 1 consists of 280 known and 219 unknown metabolites obtained from 1052 twin serum samples. Batches 2 and 3 represent additional data obtained from 5004 twin plasma samples (Table 4.1) (Suhre et al., 2011). For more details of the metabolites measured by the Metabolon platform please see Appendix H.

Table 4.1 Summary of the TwinsUK Metabolon data. Based on information from (Suhre et al., 2011)

Batch	Batch1 (serum)	Batch2 (Plasma)	Batch3 (Plasma)
N known metabolites	280	281	281
N unknown metabolites	219	175	175
Total	499	456	456

The TwinsUK Metabolon data were divided into three batches. Batch 1 consists of 280 known and 219 unknown metabolites. There are 281 known and 175 unknown metabolites for batches 2 and 3.

All metabolites in both cases and controls were residual-corrected for age and BMI. We only looked at metabolites which were measured in more than 500 individuals (non-missing values in >500 of the total set of the individuals).

4.2.4 Summary of MWAS scans of the TwinsUK Metabolon data.

Motivation

In this chapter we performed 19 different MWAS scans on our cases and controls, all testing for metabolomic differences between levels of seriousness in allergic reaction and/or differences between allergic and non-allergic cases. Table 4.2 summarises all 19 MWAS tests, including the phenotype, the model type used and the case/control selection for each test. We treated the phenotype for tests 1, 5, 6, 12, 13, 14, 15 and 19 as an ordinal trait (ordinal logistic regression applied). The phenotype for test 3 was treated as a quantitative trait

(linear mixed model applied), and the phenotype for tests 2, 4, 7, 8, 9, 10, 11, 16, 17 and 18 as a binary case/control trait (logistic regression applied).

Table 4.2 Summary of MWAS scan of the TwinsUK Metabolon data.

Test	Phenotype	Phenotype fitted as	Model type	Related twins included?	Case selection	Control selection
1	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	Yes	All	All
2	Beta-lactam allergy	Case-control	Logistic regression	No	Immediate+semi-delayed cases	All
3	Reaction seriousness	QT (1-2-3)	Linear regression	No	All	No controls
4	Beta-lactam allergy	Case-control	Linear mixed model	Yes	Immediate+semi-delayed cases	All
5	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	Yes	Immediate cases	All
6	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	Yes	Semi-delayed cases	All
7	Beta-lactam allergy	Case-control	Logistic regression	Yes	All	All
8	Beta-lactam allergy	Case-control	Logistic regression	Yes	Immediate cases	All
9	Beta-lactam allergy	Case-control	Logistic regression	Yes	Immediate+semi-delayed cases	All

10	Beta-lactam allergy	Case-control	Logistic regression	Yes	Semi-delayed cases	All
11	Beta-lactam allergy	Case-control	Logistic regression	Yes	Semi-delayed+delayed cases	All
12	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	Yes	Immediate+semi-delayed cases	All
13	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	No	All	All
14	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	No	Immediate cases	All
15	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	No	Semi-delayed+delayed cases	All
16	Beta-lactam allergy	Case-control	Logistic regression	No	All	All
17	Beta-lactam allergy	Case-control	Logistic regression	No	Immediate cases	All
18	Beta-lactam allergy	Case-control	Logistic regression	No	Semi-delayed+delayed cases	All
19	Reaction seriousness	Ordinal (0,1,2,3)	Ordered logistic regression	No	Immediate+semi-delayed cases	All

Ordered logistic regression

Many variables of interest such as diseases are ordinal and graded on a scale e.g. from least severe to most severe. When one can rank the values but the real distance between categories is unknown, then ordered logistic regression model is appropriate for the trait (Morris et al., 2010). In this case, we applied the model to allergy seriousness as an ordinal trait. In order to account for the family-based genetic structure in the twins data, we fitted ordered logistic regression models in the program STATA using the ologit function with a clustered error structure defined by the cluster() option within each family.

Logistic regression (singletons only)

In order to investigate the causes of inflation in the result of test 1, we applied logistic regression to beta-lactam allergy as a binary case-control trait restricted to immediate+semi-delayed cases, and using singleton twins only in order to remove twin correlation structure by using the glm() function in R.

Linear mixed model

The aim of test 3 was to examine whether the inflation in test results seen in the QQ plots was the result of the differences between the cases and controls or was due to some other bias. By removing the control group and only using singletons, we tried to investigate whether the inflation still appeared in our results or not. We used the EMMA software to run linear mixed models, recalculating the covariance matrix excluding the index metabolite (metabolite of interest) each time. The EMMA software is a statistical test available in the R package and is

suitable for mixed models analysis and correlation structure correction (Zhou and Stephens, 2012).

4.2.5 Post-association analysis

QQ-plot

As a post-association analysis of our MWAS scan, we created QQ plots. The quantile-quantile plot (QQ plot) is a graphical tool that compares the distribution of observed negative log p-values versus expected negative log p-values under the global null hypothesis (Ehret, 2010). This plot helps to indicate whether the study has achieved more significant results than expected under a global null hypothesis (for more information on the Q-Q plot please see chapter 3). The QQ plot also serves as a useful way to control for multiple tests within a single MWAS scan, as results are compared to a global null hypothesis across all within-MWAS tests. QQ plots do not control for multiple tests across all the different MWAS scans conducted in this chapter, so here it is important to appreciate that (1) all results presented in this chapter are preliminary; and (2) if equivalent results are seen across multiple MWAS scans, then these are unlikely to have generated under a global null hypothesis.

4.3 Results

4.3.1 Correlation structure of the metabolite data

In order to have a better understanding of the Metabolon data, we visualised correlation among the metabolite data by using a Heat-Map. A Heat-Map is a graphical tool to represent the correlation between individual/metabolites (Sugimoto et al., 2012). High correlation between individual/metabolites was represented as light yellow and low correlation between individual/metabolites was represented as a red colour (figure 4.1).

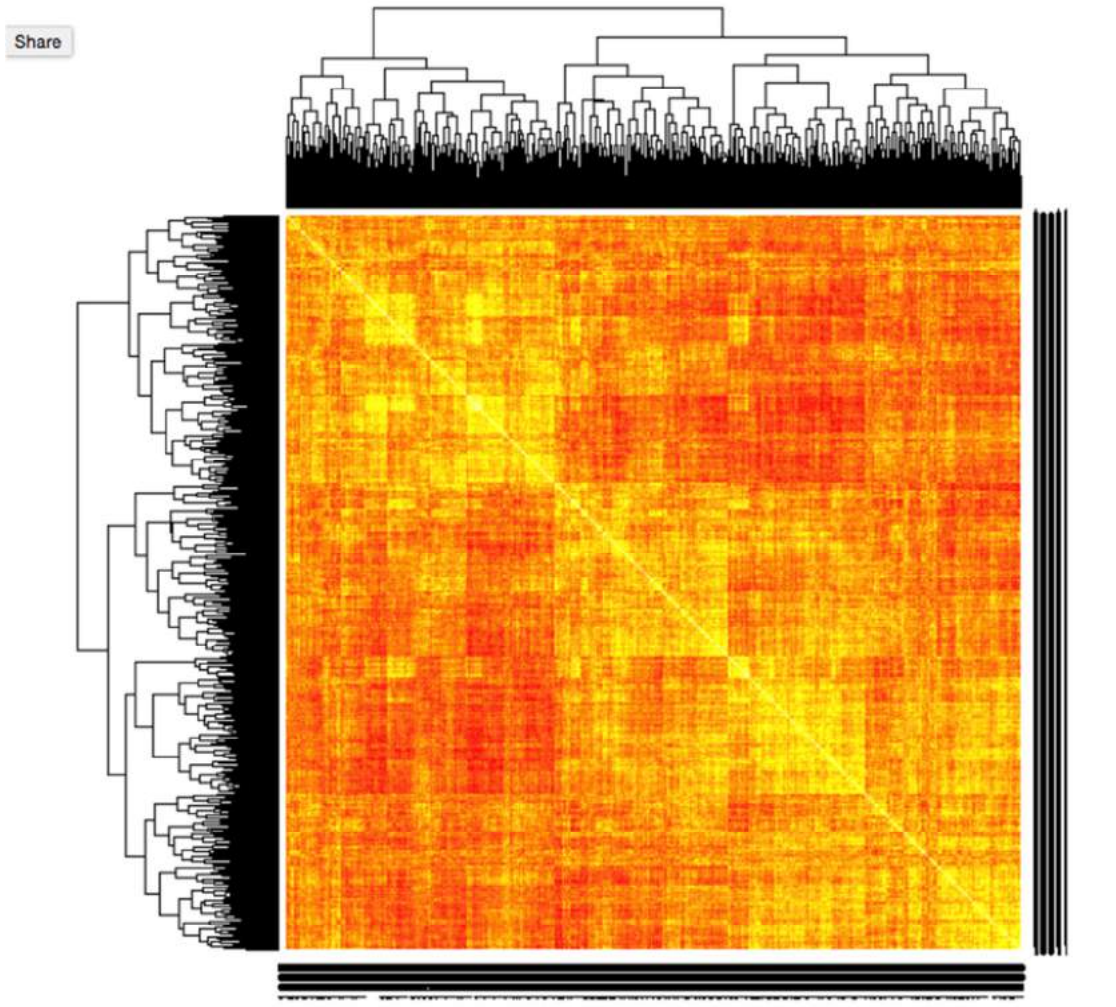


Figure 4.1 Heat-Map representation of the correlation structure of the metabolite data.
Heat-Map shows the correlation across individuals in metabolite levels (yellow = high correlation, red = low correlation). There is limited between-individual structure in these data, despite the high degree of genetic structure present in twins data.

4.3.2 Results for MWAS scans

Table 4.3 summarises the results of all 19 MWAS scans. As explained in the table below, there is an overall inflation in all Q-Q plots. The results for tests 1, 2, 3, 13 and 14 are explained in more detail in the next section. For Q-Q plots

summarizing the results of tests 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18 and 19 please see Appendix I.

Table 4.3 Results summary of the MWAS scans

Test	Hits outside QQ plot concentration band?	Sum**mary**
1	Yes	Overall metabolomic inflation in the lower end of the Q-Q plot
2	No	Overall metabolomic inflation in the lower end of the Q-Q plot
3	No	Inflation is lost in top end of the Q-Q plot, but it's still present at the bottom end
4	No	Inflation is reduced in this scan
5	No	Overall metabolomic inflation in the Q-Q plot
6	No	Overall metabolomic inflation in the Q-Q plot
7	No	Overall metabolomic inflation in the Q-Q plot
8	No	Overall metabolomic inflation in the Q-Q plot
9	No	Overall metabolomic inflation in the Q-Q plot
10	No	Overall metabolomic inflation in the Q-Q plot
11	No	Overall metabolomic inflation in the Q-Q plot

12	No	Overall metabolomic inflation in the Q-Q plot
13	Yes	Overall metabolomic inflation in the lower end of the Q-Q plot
14	No	Overall metabolomic inflation in the lower end of the Q-Q plot
15	No	Overall metabolomic inflation in the Q-Q plot
16	Yes	Overall metabolomic inflation in the lower end of the Q-Q plot
17	No	Overall metabolomic inflation in the Q-Q plot
18	No	Overall metabolomic inflation in the Q-Q plot
19	No	Overall metabolomic inflation in the Q-Q plot

4.3.2.1 Results for Test 1: MWAS scan using ordered logistic regression test on all cases and controls, treating seriousness as an ordinal trait

Q-Q plot and top hits

Figure 4.2 is the Q-Q plot of test 1 (seriousness treated as an ordinal trait). As shown in Figure 4.2 all metabolites are responding together as a correlated system. There is an overall metabolomic inflation in the lower end of the plot, and in the upper end of the plot there are 4 distinct hits lying outside the 95% concentration band of the QQ plot, which can therefore be described as

"metabolome-wide significant" following within-scan multiple test correction. As we found considerable inflation in our first test results we applied additional tests, as described in subsequent sub-sections, to investigate the source of this inflation.

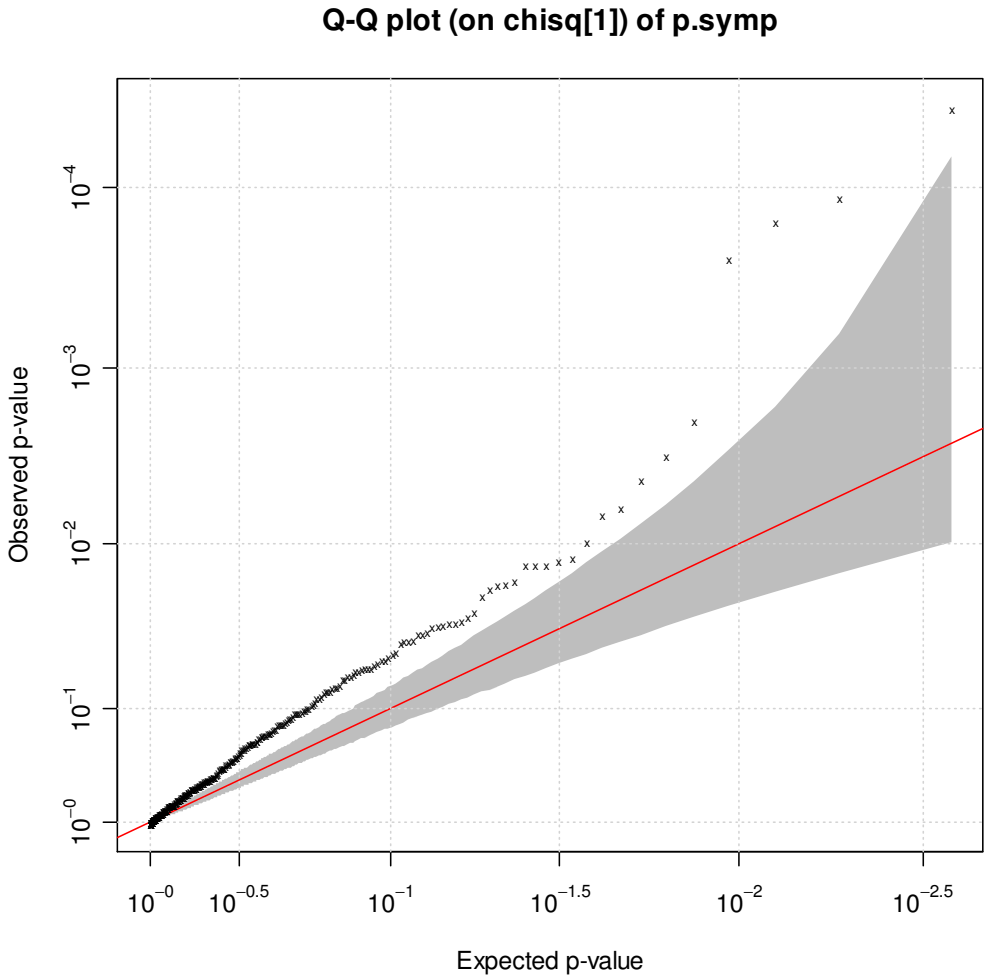


Figure 4.2 Q-Q plot for MWAS scan using ordered logistic regression on all cases and controls, treating seriousness as an ordinal trait.
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each metabolite, and the red line indicates the expected distributions under the null hypothesis. The grey shaded region represents the 95% concentration band.

Table 4.4 describes the top 4 hits for the Test 1 MWAS scan (clearly separated and lying outside the 95% concentration band of the QQ plot). We identified these metabolites by using the annotation file for the TwinsUK metabolite data. The annotation file for the metabolites consists of all known and unknown metabolites in the TwinsUK Metabolon data.

Table 4.4 Top hits for the MWAS scan using ordered logistic regression on all cases and controls, treating seriousness as an ordinal trait

ID	P-Value	Description
p_x33653	0.00043683	Unknown metabolite
p_x32735	0.00077965	Unknown metabolite
p_x33935	0.00003155	Piperine*/xenobiotics^/food component
p_x36098	0.00029946	4-vinylphenol sulfate+/xenobiotics/Benzoate metabolism

*** Piperine is the main alkaloid from black pepper**
^ Xenobiotics indicates a foreign metabolite found within body
+ 4-vinylphenol sulfate is a phenolic compound found in wine and beer
This table shows the top 4 metabolite hits associated with beta-lactam allergic reaction in the TwinsUK MWAS scan. The first two metabolites are unknown and not identified yet. The other 2 metabolites are categorized as being xenobiotics.

As shown in Table 4.4, there are two known and two unknown metabolite hits. The known metabolites are both categorised as being xenobiotics, meaning that they are foreign chemical substances that are not naturally produced in the body. Therefore, our results suggest that people who had penicillin allergy had less piperine in their system in comparison with our control group. Also people who

had experienced penicillin allergy had higher amounts of 4-vinylphenol-sulfate metabolite, compared to our control group. 4-vinylphenol-sulfate is a phenolic compound found in wine and beer.

Overall, the two known significant metabolites are categorized as being xenobiotics, which means that they are not the result of any gene expression related to beta-lactam allergic reactions. These metabolites are therefore not directly linked to any related gene. Further investigation is required to determine the function of the unknown metabolites.

4.3.2.2 Results for Test 2: logistic regression on allergic reaction as a binary case/control trait (singleton twins only)

Q-Q plot

Figure 4.3 is the Q-Q plot for allergic reaction as a binary case/control trait (immediate/semi-delayed cases singleton twins only). The result also shows an overall inflation in the test results. As I explained in the methods section, we used singleton data in this test to ensure that the inflation was not due to the use of genetically-correlated twin data.

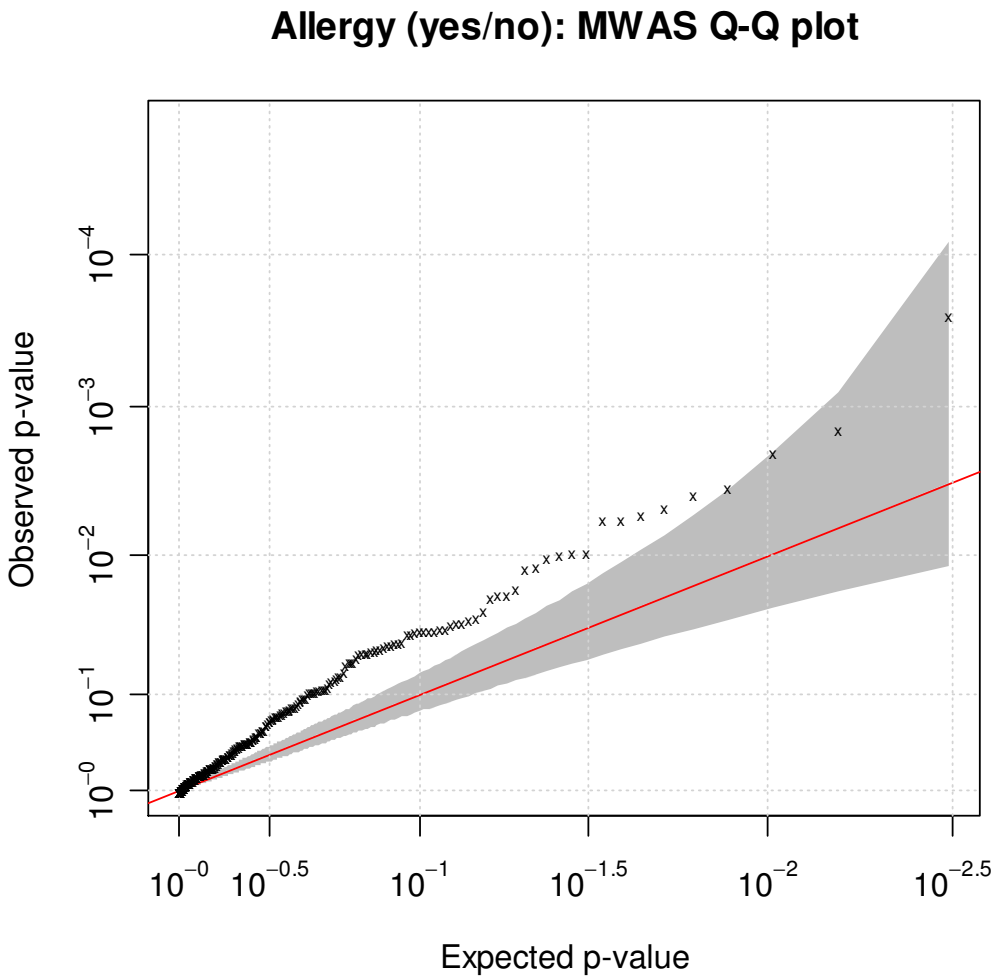


Figure 4.3 Q-Q plot for logistic regression on allergic reaction as a binary case/control trait (immediate/semi-delayed cases singleton twins only).
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each metabolite, and the red line indicates the expected distributions under the global null hypothesis. The grey shaded region represents the 95% concentration band. This Q-Q plot shows the results of an MWAS scan using singletons only, to see if the inflation seen in Test 1 is due to presence of twins genetic structure or not. There is still a general inflation seen in the Q-Q plot.

4.3.3 Results for Test 3: MWAS scan, linear regression on seriousness as a quantitative trait (singletons only and controls removed)

Q-Q plot

The aim of this test was to examine whether the inflation in test results seen in the Q-Q plots was the result of the differences between the cases and controls or was due to other bias. By removing the control group and only using singletons we tried to investigate whether the inflation still appeared in our results or not.

The Q-Q plot for Test 3 (Figure 4.4) shows no significant hits. After removing the control group the inflation is lost in the top end of the Q-Q plot, but it's still present at the bottom end. This may simply indicate that the power is too low to detect any hits when controls were removed, but still there is some inflation in the Q-Q plot.

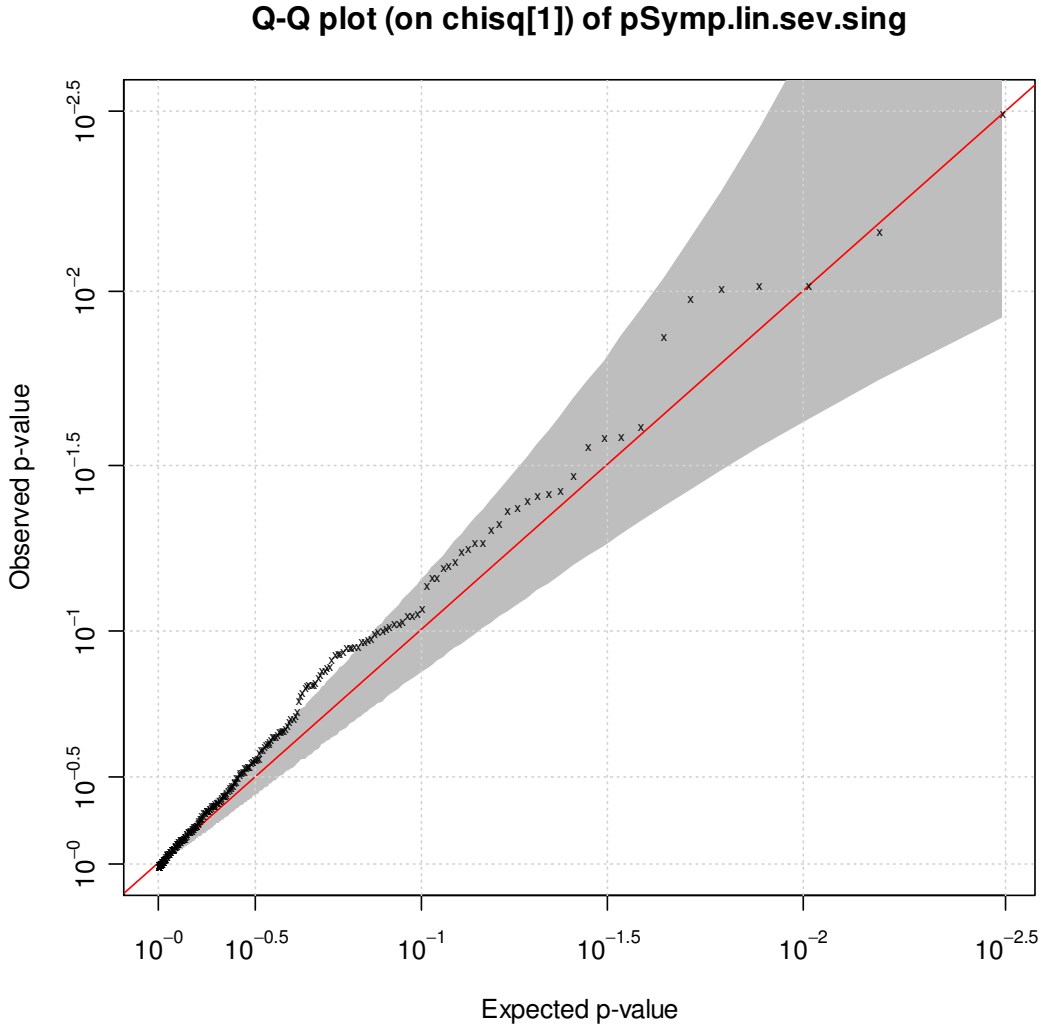


Figure 4.4 Q-Q plot for linear regression on seriousness as a quantitative trait (singletons only and controls removed).
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each metabolite, and the red line indicates the expected distributions under the null hypothesis. The grey shaded region represents the 95% concentration band. This Q-Q plot shows the results of an MWAS scan in which singletons and controls are removed. Even with no control group there is still some inflation in the lower end of the plot.

4.3.4 Results for Test 3: MWAS scan, analysis of seriousness of the reaction as an ordinal variable (singletons only)

Q-Q plot

Apart from Tests 2 and 3, the previous MWAS scans in this chapter were conducted on TwinsUK data and corrected for correlation stratification. As shown in the results section, there is an overall metabolomic inflation in our Q-Q plots. In order to see if this inflation in metabolomic expression is due to some family structure, or it is due to the reaction itself, we removed the existing co-twins from each family.

By removing the co-twins and just keeping the singletons we treated our data with standard regression methods (appropriate for when no twin structure is involved). However we applied the same QC methods to our MWAS scan. By performing these MWAS scans on singletons and comparing the result to the same tests with all the twins' involvement we were hoping to determine the cause of inflation in our results.

Figure 4.5 is a Q-Q plot for the MWAS scan on seriousness of the reaction on all cases and controls. We removed all co-twins, and thus this scan was applied to 'singletons' only. The seriousness of the reaction was analysed as an ordinal variable. The results for this test are the same as the results for Test 1. There is inflation on the lower end of the Q-Q plot, and the same 4 significant hits are seen on the upper end of the plot (see Table 4.4).

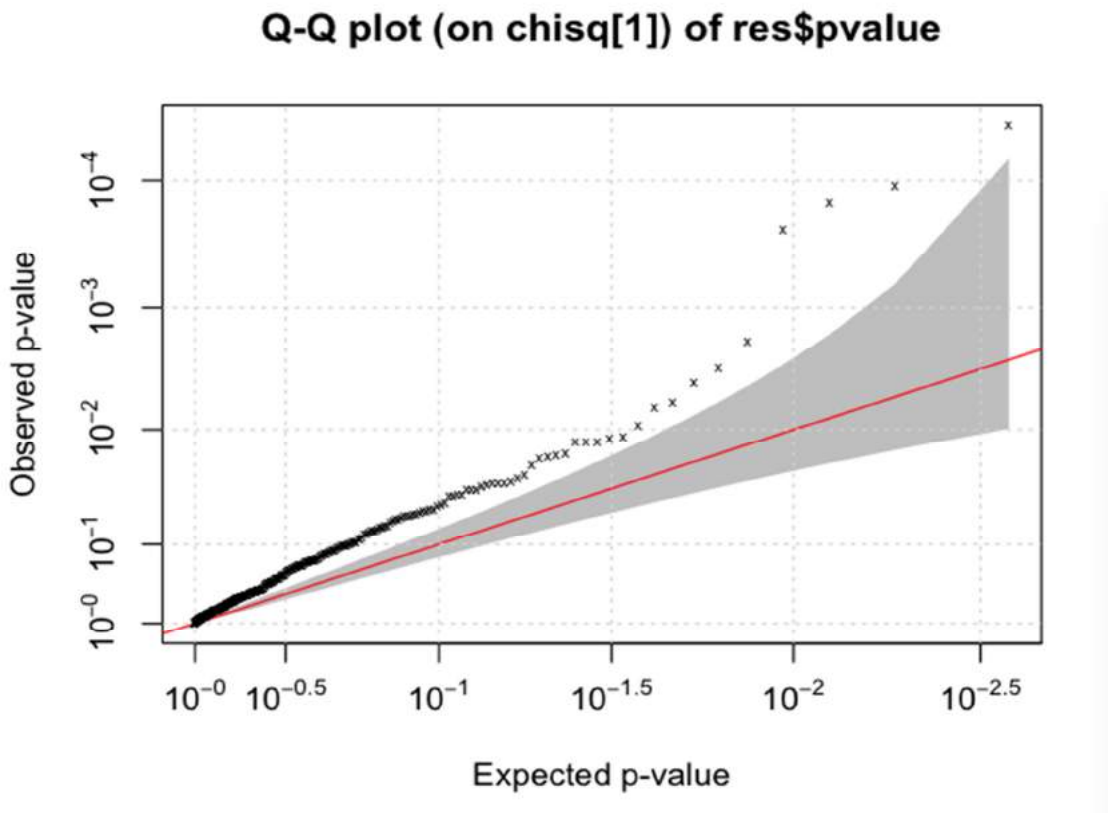


Figure 4.5 Q-Q plot for seriousness of the reaction as an ordinal variable (singletons only).
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each metabolite, and the red line indicates the expected distributions under the null hypothesis. The grey shaded region represents the 95% concentration band. Still there is an overall inflation in data. There is inflation in the lower end of the plot, and 4 significant hits are in the upper end of the plot.

4.3.5 Results for Test 4: MWAS scan, analysis of seriousness of the immediate reaction as ordinal variable (singletons only)

Q-Q plot

In this MWAS scan we removed all co-twins from the data and we performed a metabolome-wide association study on allergic cases with immediate reaction and controls. The seriousness of the reaction was treated as an ordinal variable.

Figure 4.6 is a Q-Q plot for the MWAS scan on seriousness of the reaction on immediate cases and controls. We removed all co-twins, and this scan was applied on singletons only. Seriousness of the reaction was analysed as an ordinal variable. As shown in Q-Q plot, conducting the same analysis as Test 13, just on cases with immediate reaction, we didn't detect any significant hits, likely due to the reduced power in the analysis.

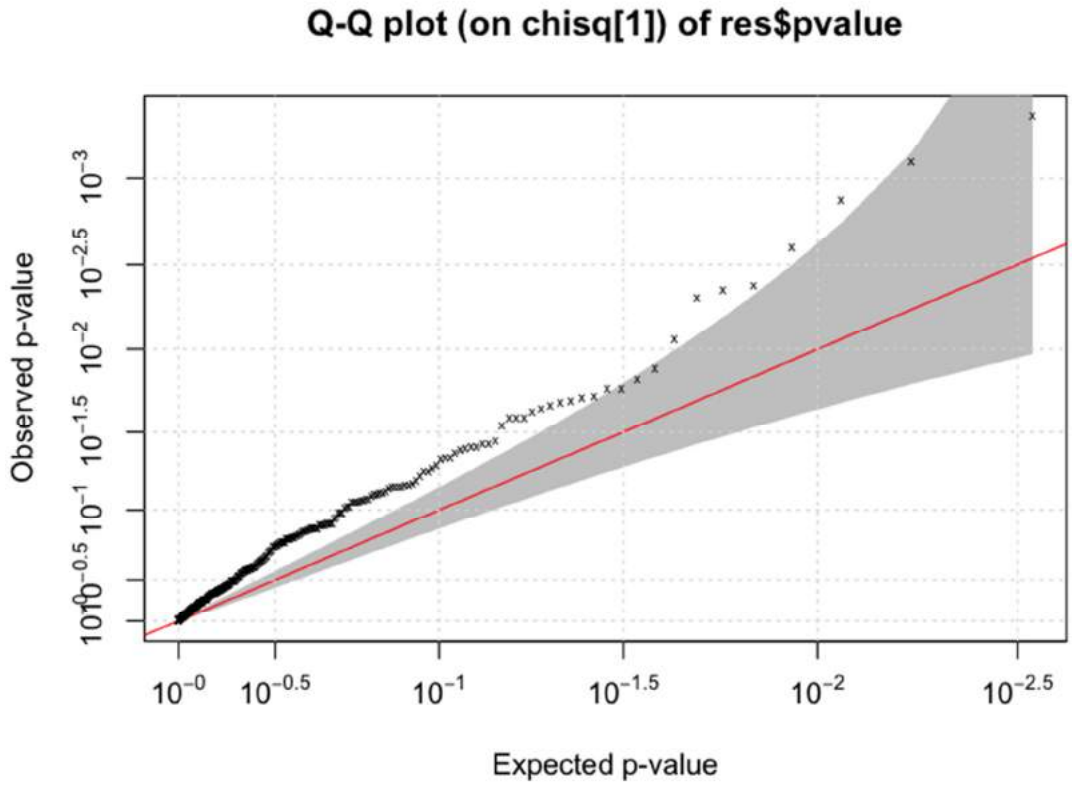


Figure 4.6 Q-Q plot for seriousness of the immediate reaction as an ordinal variable (singletons only).
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each metabolite, and the red line indicates the expected distributions under the null hypothesis. There is inflation but no significant hits.

4.4 Discussion

4.4.1 Previously published metabolomic studies

Previous studies have identified a few metabolites in human serum that result from the breakdown of beta-lactam antibiotics in the biological system. Also, a comprehensive study of genetic influences on human metabolism using twins data has revealed an atlas of genetic influences on human blood metabolites in 2014 (Shin et al., 2014). This study was conducted on 7824 adult individuals from two European population studies, including 1052 individuals from the TwinsUK cohort and 1768 individuals from the KORA (Kooperative Gesundheitsforschung in der Region Augsburg) cohort (Shin et al., 2014). They investigated 60 different biochemical pathways in their study. In this comprehensive metabolomic genome-wide association study, they discovered 84 new genetic signals affecting metabolic traits. They propose that newly identified genetically-influenced metabolites will empower future clinical and pharmacological research from better understanding of the genetic predisposition to different diseases, and will allow the identification of potential new biomarkers for drug targets and causal environmental influences on human traits (Shin et al., 2014).

Here we applied a hypothesis-free metabolome-wide association approach analysing >500 metabolites in the serum sample of 1234 individuals from the TwinsUK cohort.

4.4.2 MWAS scan on the TwinsUK cohort

In total we ran 19 MWAS scans; each time applying a different phenotype (allergy status variables) on all twins, or the same phenotype on singletons (twins without their co-twin) or on subsets of the case/control data (e.g. short time delay cases only). The aim of applying various MWAS scans was to see if there are any metabolomic differences between different types of beta-lactam allergy status or not. Out of 19 MWAS scans, 3 of them provided significant hits (defined as clearly separated hits lying outside the 95% concentration band of the QQ plot). Here I am going to briefly review the results of the key analyses in this chapter.

As a first test in this chapter, we ran an MWAS scan on all 211 allergic cases and 1023 controls. In this test we treated the seriousness of the beta-lactam allergic reaction as an ordinal variable. The result of Test 1 shows that there is a systematic inflation in the data (Figure 4.2). This test also identified 4 significant metabolite hits (Table 4.4). The first two metabolite hits are of unknown chemical structure, which means that we don't have any information regarding whether they are xenobiotics or protein by-products. The two known metabolites are piperine and 4-vinylphenol sulfate. Both the known metabolites are xenobiotics. At this stage, before any further investigation on the nature of the known metabolites, we ran more tests in order to explore the systematic inflation shown in the Q-Q plot (Figure 4.2).

As the TwinsUK data was used in this chapter, we suspected that the inflation might be due to the high level of relatedness in our data. Therefore, we ran another MWAS scan just on singletons (twins without their co-twin) in our data.

Test 2 was a logistic regression test on seriousness of the reaction, and in this test we analysed allergic reaction as a binary case/control trait (immediate/semi-delayed cases versus controls, singleton twins only). The test results show that the pattern of metabolomic differences still exists between our cases and controls, and thus it was not just due to the genetic relatedness between twins. Therefore we decided to run another scan to examine whether the inflation was the result of the metabolomic differences between allergic cases and controls or was due to some other bias.

Test 3 was a linear regression test on seriousness of the allergic reaction to beta-lactam antibiotics. In this test we analysed seriousness of the reaction as a quantitative trait. In this scan, not only did we remove all co-twins (singletons only), but we also removed all controls to allow us to focus on within-case contrasts.

There was still inflation in the results from Test 2, and Test 3. The inflation was reduced, but even when all controls were removed we still saw inflation in the data. One of the possible explanations for this systemic inflation is that all metabolites biologically respond to allergy in some correlated way; in other words, allergy can affect the whole biological system. Another possible explanation is due to batch effects as this could also affect all metabolite measurements in the same way.

Table 4.3 summarises all 19 MWAS scans conducted in this chapter. As I discussed in the results section, Test 1, Test 13, and Test 16 generated significant hits. The results were the same for all 3 tests, identifying the same set of two unknown hits and two known hits. Within-scan correction for multiple testing was performed via the application of a 95% concentration band analysis to the QQ plots. Hits lying outside this band indicate significant departure from the global null hypothesis. Hits were not corrected for between-scan multiplicity (i.e. the application of 19 separate scans), but the fact that the same set of 4 hits was returned in each case indicates that these scans were not independent, but rather were variations on the same theme and that in effect the same results were returned for each scan.

The results indicate that people who had a beta-lactam allergic reaction had less piperine in comparison to the control group. On the other hand, people who experienced beta-lactam allergic reaction appeared to have higher amounts of 4-vinylphenol sulfate metabolite in comparison to the control group. The other 16 MWAS test results were negative most probably due to the reduced number of cases and controls (and hence the power of the study was not enough to detect any positive results). There is an overall inflation in most test results, and thus metabolites appear to act as a correlated system to beta-lactam allergy.

4.4.3 Strength and limitations of MWAS on the TwinsUK beta-lactam allergy data

Here I discuss some of the strengths and limitations of the MWAS scans as performed in this chapter, which may have contributed to the results that we obtained.

Limitations of the study included: (1) a small sample size, which negatively impacts the power of a study; (2) self-reported data via questionnaire (see chapter 2 for a discussion of limitations and bias related to data collection using the self-reported questionnaire; (3) recall bias for example due to the fact that there were a large number of elderly participants in the cohort who may have found it difficult to recall the original reaction; (4) twin report bias, for example due to twins being more likely to report themselves as having a trait if their co-twin had also been diagnosed with a trait (but note this bias would only affect the scans involving co-twins); (5) **h**eterogeneity in the phenotype, for example due to combining of cases with all type of reactions (all time delays), which could introduce heterogeneity as a result of having different biological pathways amongst our cases; (6) misclassification of cases with respect to penicillin allergy (It is known that the majority of patients who consider themselves to be allergic to beta-lactam are not); (7) unknown metabolites (two top hits of the study are not characterised yet, limiting biological interpretation); (8) metabolite measurement (the TwinsUK metabolites were typically measured a long time after the beta-lactam allergy reaction, often years later); (9) power calculation (we did not run a power calculation for the TwinsUK MWAS, due to the complications on power calculations in the context of a mixed-model analysis).

Strengths of the study included: (1) the use of super controls (who reported not being allergic to anything at all); (2) diagnostic status for most cases (although we identified our case group based on self-reported questionnaire, 181(85.7%) of them were reported as being diagnosed by a doctor (see chapter 3)); (3) case and control selection from the same cohort (helps to minimize potential biases); (4) minimised technical batch effects (metabolite collection and measurement were done in the same facilities and thus the test results are less likely to be affected by possible batch effects).

4.4.4 Assessment of the known metabolite hits

As previously explained, many factors such as small sample size, unknown metabolites, time interval between reaction time and metabolite measurement, batch effects and self-reported data are limitations of this study. Furthermore, we were not able to establish any biological relationships between the two known metabolite hits and beta-lactam allergic pathways.

Taking into account the small case sample size in our study and self-reported data, as a final step, I decided to check all the positive findings in previously published studies to see if they reported the same metabolites being associated with any traits.

Although the two known significant metabolites were identified as being xenobiotics, an epigenome-wide association study (EWAS) between DNA methylation and metabolic traits (metabotypes) in human blood, conducted in 2013, identified 4-vinylphenol sulfate metabolite as being associated with CpG loci (Petersen et al., 2014). The study was conducted on 1814 participants from the KORA cohort. They measured 649 blood metabolites and looked for association with methylation at 457004 CpG sites. The study suggested that variability in 4-vinylphenol sulfate is more under environmental control than it is under genetic control (Petersen et al., 2014).

Knowing all the biases of self-reported data, and different biological pathways for immediate and delayed beta-lactam related allergic reaction, the question of whether this MWAS hit could be a false positive or not still remains. Below I compare two alternative interpretations of our findings.

The hit is a false positive

The main finding of our TwinsUK MWAS results is that there is a systemic pan-metabolite reaction to beta-lactam allergy status. Under this interpretation one could argue that the significant hits found in the MWAS scans were just down to chance, and that these findings would not be replicated in other cohorts. In other words, the results that happened to float to the top in terms of strength of association, 4-vinylphenol sulfate included, should not be given too much weight, as these metabolites happened by chance to appear at the top end of a list that was subject to considerable inflation in test statistics.

The hit is a true positive

Alternatively, we can also consider that there is a possibility that 4-vinylphenol sulfate is a genuinely strong signal that would be replicated if the study were repeated. Under this interpretation, and taking account of reports of a previously published epigenome-wide association signal involving 4-vinylphenol sulfate (Petersen et al., 2014), then our MWAS result could also have a non-genetic mechanism. In this context, having a beta-lactam allergic reaction could be interpreted as an "environmental shock" which altered the metabolome in some semi-permanent way (i.e. in a way that could still be detected years later, when the serum samples were taken). 4-vinylphenol sulfate would then be a marker for that environmental effect.

Follow-up studies

The MWAS approach is relatively new and open to improvement. In the case of our TwinsUK MWAS scan on beta-lactam allergic cases, we can propose a list of suggestions to improve future studies:

- Better phenotyped data

  Clinically proven and tested cases will improve the power of the study and will give more confidence in interpretation of the signals detected.

- Better knowledge of the metabolites

  Having more known metabolites and knowledge of their biological pathways will also improve the interpretation of future MWAS scans.

- More sensitive metabolite measurements;

  Having more sensitive metabolite measurement tools will increase the number of measured metabolites and decrease the potential influence of batch effects and other measurement related bias.

- Sample collection

  Collecting samples from different tissues will help to detect more metabolites in tissue-specific contexts.

- Time of sample collecting

  Collecting the samples right after the allergic reaction will help to detect more accurately the relevant metabolites responding in the short term to the reaction.

Finally, we note that the two unknown metabolite hits are not identified yet. Therefore part of the interpretation of this chapter depends on the identification of these two unknown metabolites in the future. Also future studies are required to validate the existence of systematic metabolomic differences between allergic and non-allergic cases.

Chapter 5: GWAS of beta-lactam allergy clinical cases

5.1 Introduction

Results from a number of candidate gene studies including an Immunochip study on a large number of immune-related genes so far suggests that genetic variability in the genes which encode for important immunological markers of an allergic response may play a role in the susceptibility to immediate allergy to beta-lactam antibiotics (see Chapter 1). These include tumour necrosis factor-α (TNF-α), Inteleukin-4 (IL-4), IL-4 receptor (IL-4R), IL-10, and IL-13 (Gueant et al., 2014, Oussalah et al., 2016).

In Chapter 3, I performed a genome-wide association study (GWAS) on self-reported TwinsUK data. The genome-wide scan yielded a single strong signal in chromosome 15. The dataset that was used in chapter 3 was quite large; however it was based on a self-reported phenotype (for further information please see chapter 3).

To complement the results of the twins GWAS, I conducted a separate GWAS (i.e. I ran these GWASs in parallel, not one after the other) on beta-lactam allergy using a different clinically-defined cohort. By running another GWAS on the same phenotype but different cohort, I aimed to replicate the results from chapter 3.

In this study, cases were recruited from the Guy's allergy clinic. The GWAS study was restricted to 50 cases who were identified as being truly allergic to beta-lactams. Power plays a significant role in GWAS studies, and most successful GWASs have required larger sample sizes (Daly, 2010a). However, there have

been pharmacogenetics GWAS studies showing a contribution of genetic variants in drug response, even with a very small sample size, which suggests that effect size of pharmacogenetic traits may be bigger than for other complex traits. For example, a study conducted in 2008 found that a common variant in SLCO1B1 was strongly associated with an increased risk of stain-induced myopathy by just using 85 cases and 90 controls (for more detail please see chapter 1) (Link et al., 2008). .

In the TwinsUK GWAS, we had a larger sample size in comparison to our clinical sample, but it was thought that by having a clinically well-defined phenotype and increasing our control size we could to some extent overcome the power issue in our GWAS study. I ran a GWAS scan on all cases using the assumption that there are shared genetic loci involved in allergic response to beta lactams, regardless of which specific beta lactam is administered. In order to look for any genetic effects that might be specific to this particular beta-lactam I ran another GWAS scan just on cases with amoxicillin allergy (''amoxicillin-only'')

5.2 Material and Methods

5.2.1 Ethics approval

Ethics approval for the study, including permission to contact the patients, was given by the NRES (National Research Ethics Service) formerly St. Thomas' Ethics Committee (reference 11/LO/0112) (See Appendix J).

During the recruitment process, written consent was obtained from participants for the use of information in medical studies (see Appendix K). During the same

process, patients were provided with a patient information leaflet, which contained information about the study and the use of their genetic data (see Appendix L). A health questionnaire was another document provided at the time of the recruitment, which contained a series of question regarding their current health status and any other existing health problems (see Appendix M).

5.2.2 Phenotype definition

In my study, the medical condition being investigated is immediate allergy to beta-lactam antibiotics. For better selection of our cases it is essential to know the specific phenotypic characteristics of the condition. For retrospective recruitment, it is also essential to know the data source and all distinguishing criteria.

The Guy's allergy clinic is the only beta-lactam allergy clinic in the southern UK. The patient records contain all available medical records such as GP's referral letters, history of all medical conditions, and treatments given. GP's referral letters to the allergy clinic typically contain:

(1) The history of allergy with all symptoms;

(2) Name of medication given;

(3) A request for clinical evaluation of drug allergy.

After evaluation, the consultant produces a letter, which contains information on the clinical history of the patient regarding the drug reaction, symptoms, and administered/used medications. These letters also include all step-by-step diagnostic tests, results and the patient's health condition before, during, and

after any tests. For our retrospective recruitment we reviewed all patients' records from 2005 to 2011, which also included all other drug allergies. These records were for patients who had followed all diagnostic steps in the Guy's clinic.

As previously mentioned after taking a detailed history from each subject, skin prick and intra-dermal tests were applied. In some cases, a specific IgE test was performed. Patients were confirmed as having allergic reaction on the basis of a positive skin test, or specific IgE. Note that having a negative skin test or negative specific IgE test was not enough to confirm patient as tolerant. As explained in chapter 1, the accuracy of the skin prick test and specific IgE test are not 100%. When the patient has a clear history of an allergic reaction but the results of skin test or specific IgE are negative, a drug provocation test is needed to confirm the allergy status.

5.2.2.1 Case selection and genotyping

Case selection

I selected cases by employing the participant inclusion/exclusion criteria written in the protocol. Cases were selected from the clinical records available in the Guy's Hospital allergy clinic.

Patient selection based on protocol

Up to 150 cases with a history of immediate allergic reaction to beta-lactams were retrospectively or prospectively recruited from the allergy clinics held at Guy's Hospital.

Inclusion criteria:

1. Age: > 18 and < 80 years.

2. History of immediate allergic reaction following administration of beta-lactams and at least one positive skin testing with one or more of the beta-lactam – derived reagents tested.

Definitions

- Immediate reaction: a history of anaphylaxis or urticaria and/or angioedema immediately following administration of beta-lactam antibiotics

- Anaphylaxis: occurs if subject presents two or more of the following symptoms: palm/sole or generalised pruritus, generalized erythema, difficulty to breathe, speak or swallow, cardiovascular symptoms (hypotension) and/or loss of consciousness.

- Urticaria: Wheals in one or several parts of the body.

- Angioedema: Localized oedema in one or several parts of the body.

- Beta-lactam skin testing includes penicillin allergenic determinants (benzylpenicillin, PPL, MDM, amoxicillin) and suspected beta-lactams.

- Positive skin testing: A wheal 3mm greater than the negative control (Lopez-de-Andres et al.).

- Atopy: Symptoms of rhinoconjuctivitis and /or asthma to common allergens (pollen, house dust mites, animals)

Exclusion criteria:

1. Age <18 or >80 years

2. Inability to comprehend or comply with the protocol

The main source of data (for case selection) was consultant letters to GPs summarising the diagnosis of patients, archived from 2005 to 2011. There were 1764 visits in total to the drug allergy clinic and to the drug provocation clinic. All visits were recorded in a patient information manager (PIM) tool, and could be sorted by: date of clinic, consultant's name and clinic name.

Using the PIM, I extracted hospital/NHS numbers for all patients who visited the beta-lactam allergy clinic from 2005 to 2011. Each record was quickly scanned for any keywords that served as clues of allergy to beta-lactam antibiotics or a test for one. All records with these keywords were printed. Printed medical records were reviewed and records of interest (cases) were identified who were confirmed as having immediate (type I) allergy to beta-lactam antibiotics (i.e.

positive skin prick test, positive specific IgE test or positive DPT result). After identifying all possible cases, required information such as address, phone number, GP's contact details, date of birth, age, gender, and ethnicity were extracted using the PIM.

A total number of 1431 medical records were reviewed. Out of these records, 145 individuals were identified as suffering from immediate allergy to one or more beta-lactam antibiotics and 7 cases of delayed allergy to beta-lactam antibiotics (Table 5.1). A total of 262 patients were identified as beta-lactam tolerant and 190 individuals did not complete all the routine diagnostics for the clinician to make the appropriate decision. The remainder of records was cases of allergy to other drugs unrelated to beta-lactam antibiotics. Out of 145 allergic cases, I identified 31 duplicates, thus I retrospectively identified a total of 104 cases of immediate allergy to beta-lactam antibiotics.

Table 5.1 Drug allergy clinic review from 2005 to 2011

	Drug allergy clinic from 2005 to 2011
Total Number of patients who attended the allergy clinic	1431
Patients with immediate allergy to beta-lactam antibiotics	104 (7%)
Patients with delayed allergy to beta-lactam antibiotics	7 (1%)
Not allergic to beta-lactam antibiotics	262 (18%)
Incomplete diagnosis	190 (13%)
Other drugs	868 (61%)

This table shows summarises all available patients records in the Guy's allergy clinic from 2005 to 2011. Only 104 patients out of 1431 were confirmed with immediate allergy to beta-lactam antibiotics.

Identified cases were asked to attend the clinic in order to donate blood for our study, or else were sent postal blood kits and instructed to return them to us through their GPs for DNA extraction. Each case was asked to fill in a simple health questionnaire (see Appendix O) and consent form (see Appendix M) as well as donating blood. We did not include 52 patients due to lack of response to our letters and calls, 2 patients due to unwillingness to provide a blood sample and 2 patients who were black African. A total of 48 cases were recruited for the study (please see Table 5.2).

Case genotyping

I extracted DNA from whole frozen EDTA-treated blood, using the Nucleon Genomic DNA Extraction Kit BACC3. DNA extraction was performed according to the Standard Operating Procedure of the Genomic Laboratory of the Department of Twin Research & Genetic Epidemiology at KCL.

After DNA extraction, these individuals were genotyped on the Illumina Omni2.5 platform at the BRC Genomics Core facility (NIHR Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London, 16th floor, Tower Wing, Guy's Hospital, Great Maze Pond, London SE1 9RT).

Table 5.2 Detailed clinical information on the 48 GWAS cases recruited from the Guy's allergy clinic

ID	Sex	Age	Existing Medical Conditions	Current Medication	Co morbid allergic Conditions	Atopy	Known Allergies	Culprit Drug	Type of reaction	Name of Test	Positive Test Results
1	F	26	13	16	NA	0	NA	PEN	IM	SKIN	BP
2	F	51	8,19	3,13,18	2,3	1	2,6	PEN	IM	SKIN	NA
3	F	77	1,2	18	NA	0	NA	PEN	IM	SKIN	MDM
4	F	73	1,3	9,14	NA	0	NA	PEN	IM	SKIN	AX, AUG
5	F	61	3,5,10,12	9,10,19,22	NA	0	4,6	PEN	IM	SKIN	BP
6	F	37	8,22	3,6	2,4	1	2	AUG	IM	SKIN	AUG
7	F	58	3,5,6,8,10,12, 13,17,19	1,3,6,12,13,17, 19	1,2,3	1	6	CEFU	IM	SKIN	CEFU, CEFO
8	F	54	5,16,18	1,10,22	3	1	4,5,6	CEP	IM	SKIN	AX
9	M	57	3,5,14,15	7,10,13,19,22	NA	0	2	PEN	IM	SKIN	PPL
10	F	52	2,3,5,18,19	1,10,12,19	3	1	2	PEN	IM	HI	PEN
11	F	64	3,5,8	3,9,10,19	NA	1	2	AX	IM, DE	SKIN	PPL
12	F	39	9,2	10,12,13,17,18	NA	0	6,7	PEN	IM	SKIN	NA
13	F	67	6,22	14,15	NA	1	7	CEF	IM	SKIN	CEFO, CEFTR
14	F	62	2,3,5,6,8,13,18 ,19,20	3,6,12,13,15,16 ,18,19	2,3	1	8	PEN	IM	SKIN	BP
15	F	64	19	1,8,12	3	1	6	AX	IM	SKIN	AX
16	F	42	8,17,19	NA	1,2,3	1	2,3,6,8	AX	IM	SKIN	AX
17	M	66	NA	6,10	NA	0	6	AX	IM	RA	0
18	F	48	NA	NA	NA	1	NA	AX	IM	SKIN	AX
19	F	47	NA	14	NA	1	NA	CEF	IM	SKIN	PPL, CEFU
20	F	63	6,16	15	NA	0	6	AUG	IM	SKIN	BP, AX
21	F	76	5,8,17	14,18	1,2	1	2,3,7,8	PEN	IM	SKIN	PPL
22	F	51	6,8,10,13,17, 18,19,20,22	1,3,5,6,11,13, 15,16,18,22	1,2	1	1,2,3,6,8	PEN	IM	SKIN	BP
23	F	36	NA	NA	NA	1	4,6	PEN	IM	SKIN	NA
24	F	45	NA	14	NA	0	2,8	AX	IM	SKIN	AX
25	F	73	3,5	10,12,19	2	1	6,7	PEN	IM	SKIN	FLUC
26	M	54	NA	19	NA	1	NA	CEP	IM	SKIN	CEFU, CEFO, CEFT, CEFTR
27	F	69	22	10,12	NA	0	2,3	FLU	IM	SKIN	FLUC
28	M	62	5	1,10	NA	0	NA	PEN	IM	SKIN	AX, MDM
29	F	55	3,8,17,18,19	1,3,13,15,22	1,2,3	1	2,3,4,5,6,8	PEN	IM,DE	SKIN	BP, CEFO
30	F	66	18	5	NA	0	1,6	PEN	IM	DPT	PEN
31	F	57	NA	1	NA	0	NA	PEN	IM	SKIN	BP
32	M	71	NA	NA	NA	0	NA	PEN	IM	SKIN	BP

33	F	33	8,17	NA	1,2	1	6	PEN	IM	SKIN	BP
34	F	38	17,20,21	4,8	1,4	1	NA	AX,CEP	IM	SKIN	AX
35	M	74	3,8,22	3,6,10	NA	0	NA	AX	IM	SKIN	AX
36	M	64	2,3,5,8	1,3,6,10,12,13, 19	2	1	6,8	AX	IM	SKIN	PPL, MDM, AX, CEP
37	F	54	18	8	NA	1	NA	AX	IM	SKIN	AX
38	M	52	19	6,10,13,22	3	1	NA	FLU	IM	SKIN	FLUC
39	F	33	8	3	2	1	NA	AX	IM	RA, HI	+RA
40	F	61	19	10	3	1	2	AX	IM	SKIN	AX
41	F	68	2,3,5,11	10,19	NA	0	NA	CEFU	IM	SKIN, RA	AX
42	M	75	4,6,710,12	NA	NA	0	6	PEN	IM	SKIN	PPL
43	F	58	17,19,22	1,8,9	1,3	1	2,6,8	AUG, PEN	IM	SKIN	PPL
44	F	28	8,9,10,13,17, 21	3,13,16	1,2,4	1	6	PEN	IM	DPT	PEN
45	M	66	3,8,13,19	1,16,19	NA	1	2,6	AUG	IM	SKIN	PPL
46	M	78	22	6,9,22	NA	0	6	PEN	IM	SKIN	BP, PENG
47	M	63	NA	NA	NA	0	NA	AX	IM	SKIN	AX
48	F	58	8,19,22	3	2,3	1	NA	AX	IM	SKIN	MDM, BP

Existing Medical Conditions: 1, Cancer; 2, Cardiovascular Disease; 3, High Blood Pressure; 4, Alcohol/Drug Abuse; 5, High Cholesterol; 6, Lung/Respiratory Disease; 7, Infectious Disease; 8, Asthma; 9, Immune Disorders; 10, Obesity; 11, Stroke; 12, Diabetes; 13, Depression; 14, Liver Disease; 15, Kidney Disease; 16, Neurological Disorders; 17, Eczema; 18, Menopause; 19, Hay Fever; 20, Contact Dermatitis; 21, Urticaria; 22, Other

Current Medication: 1, Vitamins/Minerals; 2, NSAIDs; 3, Asthma Medication; 4, Oral Contraceptives; 5, Sedatives/Sleep Aids; 6, Prescription Pain Medication; 7, Oral Hypoglycemics; 8, Hormones; 9, Diuretics; 10, Statins; 11, Herbs; 12, Aspirin; 13, Antihistamines; 14, Thyroxin; 15, Steroids (Nasal/Topical); 16, Antidepressants; 17, Insulin; 18, Antibiotics; 19, Other Blood Pressure Tablets; 20, Anticoagulants; 21, Antifungals

Co morbid allergic Conditions: 1, Eczema; 2, Asthma; 3, Hayfever; 4, Urticaria

Atopy* was defined based on a personal history of atopy/or clinical records

Atopy: Atopic case If the cases have one or more of the following conditions - Eczema, Asthma, Hayfever, Urticaria.
Non Atopic definition If cases do not have any of the following conditions - Eczema, Asthma, Hayfever, Urticaria.
1, Atopic; 0, Non Atopic

Known Allergies: 1, Nickel/Metals; 2, Flowers/Trees/Grasses; 3, Fragrance; 4, Latex; 5, Rubber; 6, Medicines; 7, Insects; 8, Animals

Culprit Drug: Culprit Drug Definition = The actual drug that causes the allergic reaction at the first place
AX, Amoxicillin; PENI, Unrecalled penicillin; AMP, Ampicillin; AUG, Augmentin; CEP, Cephalosporins; CEF, Cefuroxime; FLUC, Flucloxacillin

Type of reaction: IM, Immediate reaction; DE, Delayed reaction

Name of Test: SKIN, Skin test; RA, RAST test; DPD, Drug Provocation Test/Challenge Test; HI, History/NO test strong supporting history of the reaction

Positive Test Results: PPL, penicilloy-polylysine; MDM, minor determinants BP benzyl-penicilloyl; AX, Amoxicillin; PEN, Penicillin; PENG, Unrecalled penicillin; AMP, Ampicillin; AUG, Augmentin; CEP, Cephalosporins; CEF, Cefuroxime; FLUC, Flucloxacillin; CEFU, Cefuroxime; CEFO, Cefotaxime; CEFT, Ceftazidime; CEFTR, Ceftriaxone

This table contains information about the current health status, any existing atopic conditions of the patients, comprehensive information of their beta-lactam allergic reactions, and the results of the allergic test that was performed by the clinicians. It also contains Information regarding current health status extracted from the health questionnaire, which was filled in by the participants at the time of recruitment. The table above also contains information about the participants' age and gender. The first column gives information of any existing medical conditions.
The 'current medications' section gives a list of medications that participants were currently using and information on other existing allergic conditions. In the 'atopic' column, if any of the participants had one or more co-morbid allergic conditions, they are listed as being atopic, otherwise they are listed as non-atopic. In the 'known allergies' section, we questioned the patients for any known allergic reactions to any of the listed substances. The last four columns provide information on their beta-lactam allergic reaction. The 'culprit drug; section gives the name of the actual beta-lactam that caused the allergic reaction in the first place. The 'type of reaction' section is categorised into immediate and delayed reactions, based on the interval between taking the drug and showing the first signs of allergic reaction. The next column describes the type of the test which was used to confirm and identify the patients' allergic reaction. The final column displays the positive test results for any of the tested beta-lactams.
Following recruitment the participant, a detailed questionnaire was completed by each participant. The questionnaire contained information which was extracted to complete the table. The beta-lactam allergic reaction was extracted from the patients' records as previously described.

5.2.2.2 Control selection and genotyping

Our original intention was to recruit up to 150 matched, beta-lactam tolerant controls from the allergy clinic at Guy's Hospital. Our original inclusion criteria were age (> 18 and <80 years); no clinical history of immediate allergy reaction to beta-lactam antibiotics; and negative skin testing to beta-lactam antibiotics and tolerant of beta-lactam antibiotics. Our exclusion criteria were age (<18 or >80 years); immediate reaction to beta-lactam antibiotics; and inability to comprehend or comply with the protocol.

However, the number of selected cases didn't reach the protocol requirement and therefore in order to increase the study power we decided to use a larger control group than what was planned at the beginning. Having a large selected control group required more recruitment time and cost, which was out of our study budget. Taking account of the required time and money for control selection, and also taking account of the relatively low prevalence of beta-lactam allergy among patients taking beta-lactams, we decided to use the WTTCC2 control data (a publically available set of UK population controls). The WTCCC2 data includes 2,737 healthy blood donors from the UKBS (United Kingdom Blood Service) collection and 2,930 individuals from the 58C (1958 Birth Cohort) dataset. These individuals were genotyped on both the Illumina 1M and Affymetrix 6.0 genotyping arrays, as part of the Wellcome Trust Case-Control Consortium 2 study. 5,175 individuals from the WTCCC2 dataset passed our quality control filters (following the same protocol as described in Genetic Analysis of Psoriasis et al., 2010).

5.2.3 Quality control for cases

I checked all 48 cases for gender mismatches, and found no mismatched individuals. Thus we did not eliminate any individuals in this stage. After filtering, we had 13 males and 35 females.

After applying minor allele frequency (MAF)<0.015 and SNP missingness (GENO)>0.025 filters to our data, 74789 SNPs failed the missingness test (GENO>0.025), also 903293 SNPs failed the frequency test (MAF<0.015). After frequency and genotyping pruning, there were 1427979 SNPs left.

I used a threshold of p<=1e-06 for the test of departure of SNP genotype frequencies from Hardy-Weinberg Equilibrium (HWE). 3 markers were excluded based on the HWE test (p<=1e-06) and a total of 1427976 SNPs were left.

To screen out poor quality DNA samples, I used the --mind command in PLINK 'mind' sets the maximum rate of per-individual missingness. I used 'mind'>0.03 (this allows individuals with <3% missing SNP data to pass). No individuals were removed for low genotyping (MIND>0.03).

5.2.4 Population structure check – case data merged with HapMap3 data – performed using MDS

In GWAS analysis, there is always a risk of generating a false positive result that arises from genetic differences between different ancestry backgrounds, rather than genes associating with disease. This is particularly a concern when case and

control data are combined from different sources, as they were here (Weale, 2010).

In order to overcome this issue in our analysis, I used multidimensional-scaling (MDS) as a method to detect population stratification (PS). The MDS method visualises observed genetic distance among individuals. MDS analysis was performed by using PLINK.

Before merging our case data with our control data (WTCCC2), I checked for any ancestry issues in our case data. We restricted our data to SNPs that overlapped with the control data (WTCCC2). The steps involved in this process are outlined below.

5.2.4.1 Determination of overlap of SNPs with WTCCC2

In this stage, in order to determine overlap set of SNPs, I sorted the .bim files for cases and controls (Illumina 1M SNPs only), then by using "join" command I arrived at the common set of SNPs on both datasets. There are 463962 SNPs that intersect (overlapped).

5.2.4.2 Creation of a merged Case + HapMap Phase 3 dataset

The HapMap 3 dataset represent populations with large communities in the UK (i.e. European (CEU), African (YRI), Asian (JPT+CHB) and Indian (GIH)) (Duan et al., 2008). By using the --bmerge command, I merged our case data with the HapMap3 data and created a new file. The HapMap data contained 257 males, 274 females.48 cases and 112 CEU, 113 YRI, 86 JPT, 84 CHB, 88 GIH controls.

5.2.4.3 Application of MDS to the new merged dataset

I applied MDS to the merged dataset and checked the first two MDS axes (MDS1 vs MDS2) and also MDS1 vs MDS3. As can be seen in both figures, all cases cluster with CEU. In figure 5.1 and 5.2,the black colour represents our cases, the red colour represents CEU, green colour represents YPI, blue colour represents (JPT+CHB) and cyan colour represents GIH.

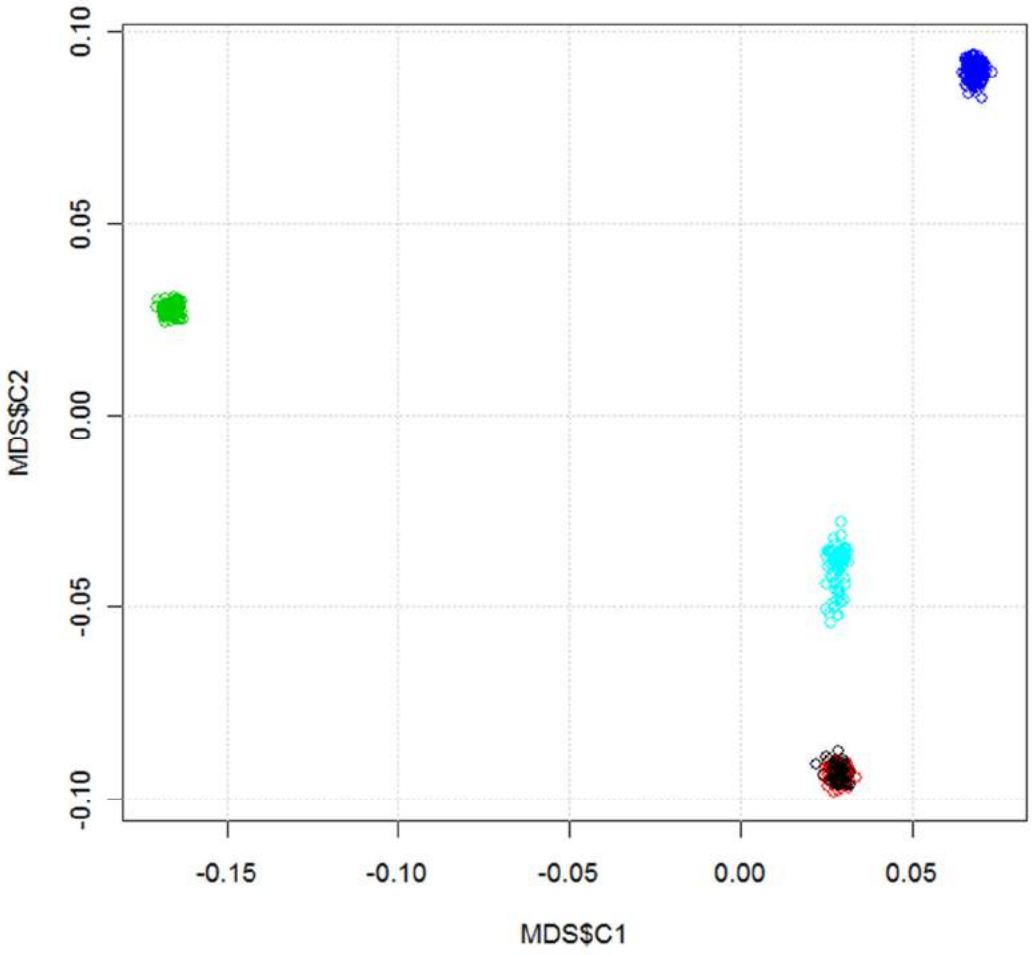


Figure 5.1 Results of MDS analysis for major ancestral population clustering.
In this figure the black colour represents our cases, the CEU population is marked as red, the green colour represents the YPI population, the cyan colour stands for the GIH population and finally the blue colour represents the East Asian (JPT+CHB) population. All cases cluster with CEU, which indicates that all our cases are likely of European ancestry.

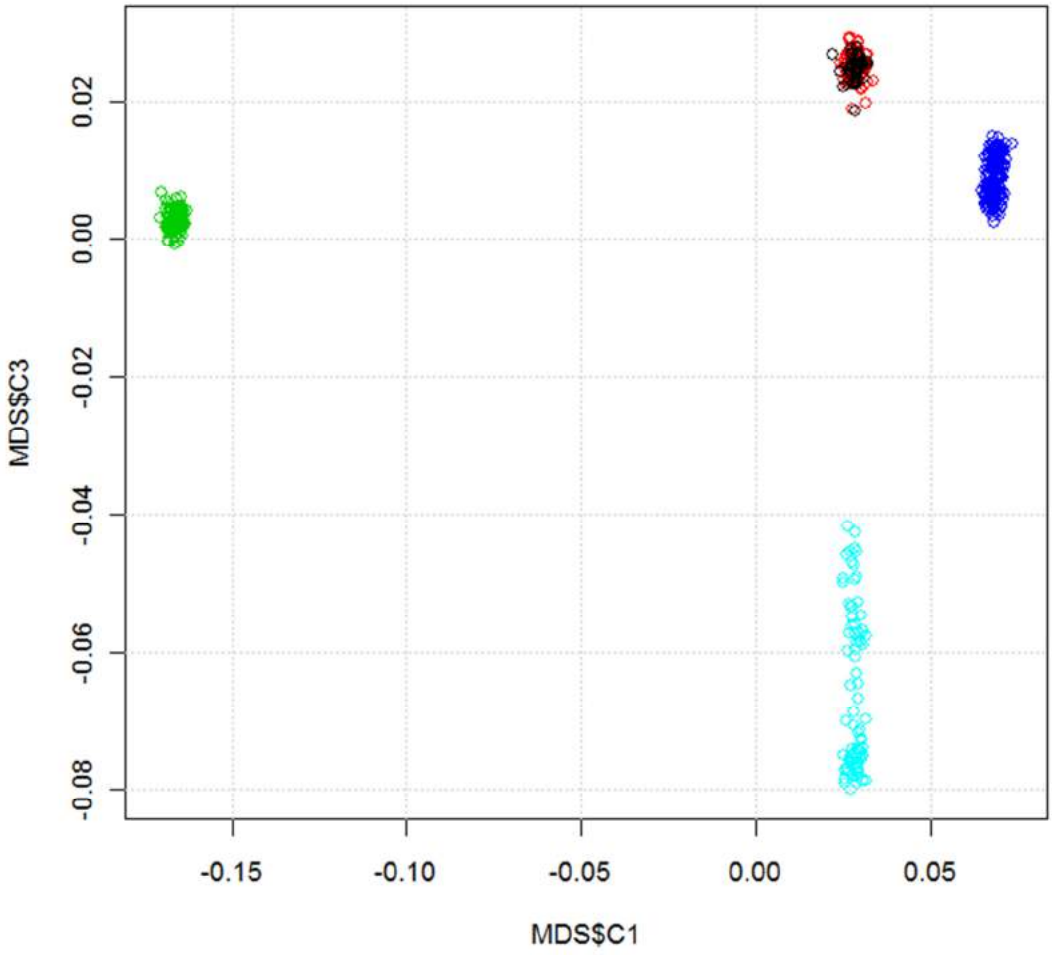


Figure 5.2 Results of MDS analysis for major ancestral population clustering.
In this figure the black colour represents our cases, the CEU population is marked as red, the green colour represents the YPI population the cyan colour stands for the GIH population and finally the blue colour represents the East Asian (JPT+CHB) population. All cases cluster with CEU, which indicates that all our cases are likely of European ancestry.

As a result of the population structure check, no individuals was removed. All individuals appeared to be of European ancestry, and all had reasonably complete genotype information.

5.2.5 Quality control and merging of control data

All the QC steps for our control dataset were applied to an 'intersect' control dataset based on a set of overlap SNPs (see section 5.2.3). First I created a reduced 1958 Birth Cohort (58C) dataset by using the -extract command, then by using the same command I created a reduced version of the National Blood Service (NBS) dataset as well. The new datasets were called 58Csmall and NBSsmall (as both datasets were smaller than the original datasets). By using the –bmerge command, I then merged the two 'small' datasets into a single version of our control data (WTCCC2).

I then applied the same QC steps to our control data set that I had used for the case data (see case QC section). Finally, by using the –bmerge command, I merged the final case dataset and final control dataset together for further GWAS analysis.

5.2.6 Power Calculations

Because of the small case sample size, before running the GWAS scan, we needed to determine what effect sizes we could pick up in the GWAS scan with 48 cases and 5139 controls. Statistical power calculation is an analysis to calculate the probability that the test can correctly reject the null hypothesis (H_0) and accept the alternative hypothesis (H_1).

To calculate our study power we used the Quanto program (http://biostats.usc.edu/Quanto.html), which can calculate required sample size and power for genetic studies. Assuming a log-additive genetic model and a

population prevalence of 0.01, and setting a genome-wide significance threshold of 5e-8, for a common SNP (MAF=0.5) we found our study had 80% power to pick up a Genotypic Odds Ratio of 4.3 and 90% power to pick up a Genotypic Odds Ratio of 4.9, while for a rare SNP (MAF=0.05) we found our study had 80% power to pick up a Genotypic Odds Ration of 6.3, and 90% power to pick up a Genotypic Odds Ratio of 7.0. We therefore determined that our study was only powered to detect very strong GWAS signals.

5.2.7 GWAS scan on all cases

We started by running a simple GWAS scan without any covariates. We applied logistic regression using SNP alleles coded (0,1,2) for each genotype ("1"=heterozygotes) – i.e. a log-additive trend model, and no covariates for population structure were added.

By not using any covariates, it was possible that background inflation in our QQ plots and Manhattan plots would be generated. However, as shown in the next section, there was in fact no evidence for genomic inflation. We therefore proceeded with this model for GWAS scan on all cases. By using all cases, this scan assumes that common genetic loci underline the genetics of beta-lactam allergy, regardless of the specific beta-lactam being administered.

5.2.8 GWAS Fisher's Exact Test (FET) scan restricted to amoxicillin cases only

The GWAS scan on all allergic cases in our clinical cases showed no positive findings (see results section).

Therefore we decided to run another GWAS scan on just amoxicillin allergy cases. The reasoning behind this was that a less heterogeneous phenotype might yield hits that were masked by the previous analysis. This GWAS scan was performed on 11 amoxicillin allergic cases. Due to the small sample size in this test we decided to use Fisher's Exact Test to achieve this scan. Fisher Exact Test is a statistical exact test applicable to all sample sizes, including small sample size. This test is well-behaved in low-n situations, and thus one can calculate the exact significance of the deviation from a null hypothesis even with a small sample size (Lin et al., 2015).

5.3 Results

5.3.1 Characteristics of the 48 GWAS cases

Table 5.3 summarises some characteristics of the 48 GWAS cases recruited from the Guy's Hospital allergy clinic from 2005 to 2011. More complete information is provided in Table 5.2.

Table 5.3 Characteristics of the GWAS cases

Name of register	Guy's Hospital allergy clinic patients from 2005-2011
Country	United Kingdom
Kind of ascertainment	Recruited retrospectively and prospectively
Mean age	56
Sex (male/female)	(13 (27%)/35 (73%))
Allergy to particular drug	PEN (45.83%), AUG (6.25%), CEFU (4.16%), CEP (4.16%), AX (27.08%), CEF (4.16%), FLU (4.16%), AUG/PEN (2.8%), AX/CEP (2.8%)
Number of immediate/delayed/immediate and delayed cases	Immediate 46 (95.8%)/delayed 0/immediate and delayed 2 (4.2%)
Atopic/non atopic	29(60.4%)/19(39.6%)

AX, Amoxicillin; PEN, Penicillin; AUG, Augmentin; CEFU, Cefuroxime; CEP Cephalosporins, CEF, Cefuroxime

5.3.2 Q-Q plot and Manhattan plot for all cases

Figure 5.3 shows a QQ plot for our GWAS scan on all cases. A QQ plot displays the observed p-values for each SNP, ordered from highest to lowest on a –log10 scale, against expected p-values under a global null hypothesis (i.e. the p-values expected if there were no true association signals in the data (Ehret, 2010)).

Since the points adhere to the 1:1 line in the lower left portion of the plot, this indicates there is no evidence for genomic inflation affecting all SNPs. The upper end of the plot shows some "hits" – i.e. association signals that can't be explained under the global null hypothesis. However, these "hits" could still be false positives, as I explain below.

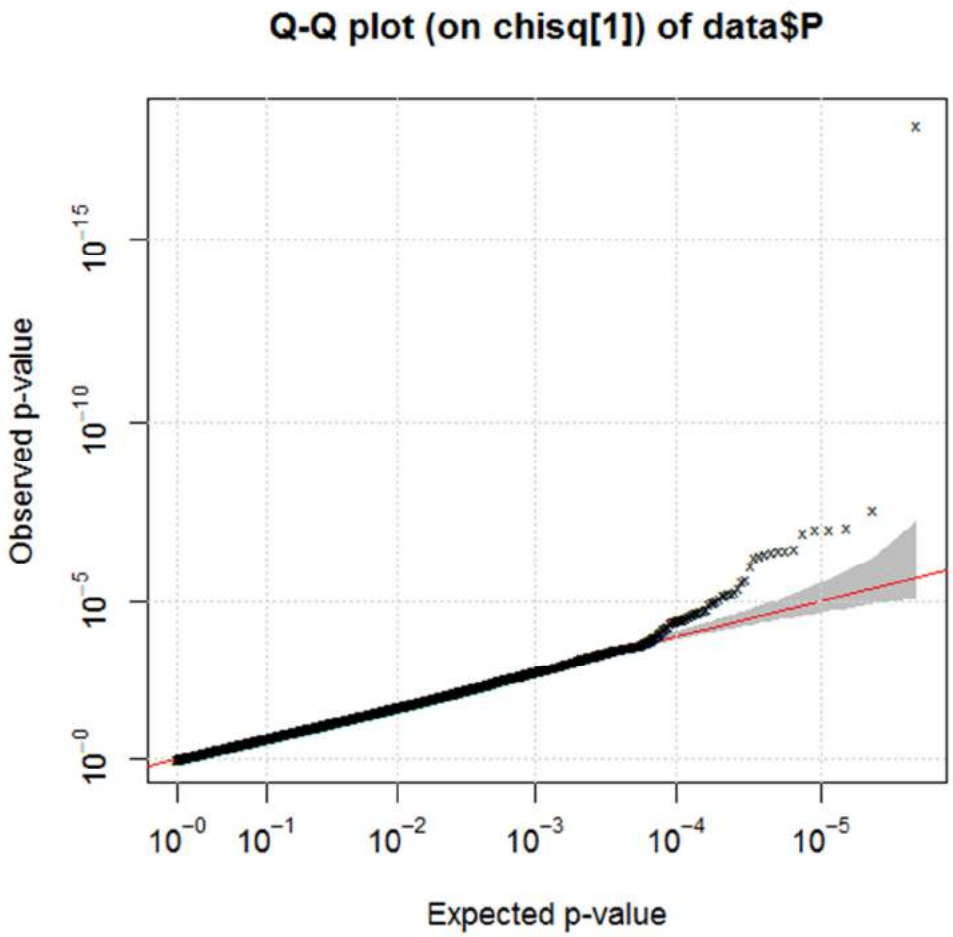


Figure 5.3 QQ plot for the GWAS scan on all cases.
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each SNP. The red line indicates the expected distribution under the global null hypothesis. The grey shaded region represents the 95% concentration band. As can be seen in this plot, there is no genomic inflation in lower end of the plot, while some apparent "hits" appear at the upper end of the plot.

In running the GWAS scan on all cases, we did not explicitly correct for population structure, but even so the QQ plot was free of any evidence of inflation. If we had seen any inflation in the plot, we would have introduced population structure covariates, such as PC or MDS axes, into the analysis to correct the inflation (Weale, 2010). But in the event, we found that the GWAS scan on all cases did not show any genomic inflation. Genomic inflation is a

problem that increases with sample size so one reason for the lack of inflation may simply be the very small number of cases in the analysis. Figure 5.4 provides a Manhattan plot for the GWAS scan on all cases.

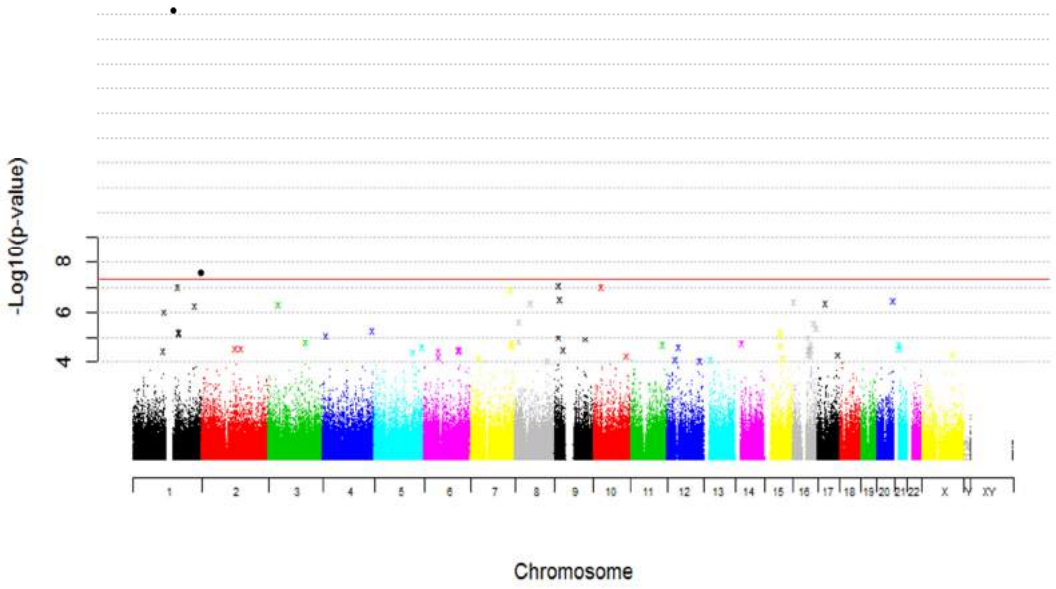


Figure 5.4 Manhattan plot for the GWAS scan on all cases.
The x-axis shows the chromosomal position and the y-axis shows the –log10 p-value. Chromosomes are shown in alternate colours. The red horizontal line indicates a threshold of genome-wide significance at a p-value of p<5e-8. Each point represents a p-value for the SNP association test. There are two significant hits in this plot, but they appear sporadic with no support from local SNPs in LD.

5.3.2.1 Top hits and LocusZoom plots

Table 5.4 lists the top 30 hits from the GWAS scan on all cases, ordered by p-value.

Table 5.4 The top 30 hits for the GWAS scan on all cases. Hits are listed from lowest to highest p-value.

CHR	SNP	BP	A1	TEST	NMISS	OR	P
1	rs6657114	147377718	A	ADD	5176	12.32	7.008e-19
1	rs12404733	247690417	C	ADD	5133	0.1538	2.684e-08
9	rs10810585	16671045	C	ADD	5160	0.2105	8.367e-08
10	rs11008285	31130983	C	ADD	5160	0.1541	9.245e-08
1	rs10458392	162092110	C	ADD	5093	0.2254	9.316e-08
7	rs10260314	148280705	C	ADD	5137	0.1219	1.14e-07
9	rs7871217	18929692	C	ADD	5164	0.2265	3.067e-07
20	rs4925203	60451598	A	ADD	5160	0.2386	3.417e-07
16	rs9925071	5784107	G	ADD	5184	3.176	3.728e-07
8	rs939341	61032713	C	ADD	5158	0.1685	4.007e-07
17	rs3764421	29167653	C	ADD	5117	0.152	4.312e-07
3	rs10865893	40113022	C	ADD	5155	0.2809	4.973e-07
1	rs2840945	225859141	C	ADD	5114	0.2662	5.515e-07
1	rs8453	115259599	A	ADD	5177	0.2383	9.754e-07
8	rs9325854	19151139	A	ADD	5182	4.317	2.454e-06
16	rs4556797	80037987	G	ADD	5127	0.2292	2.77e-06
16	rs6540308	86554024	C	ADD	5139	0.28	4.201e-06
4	rs2278924	183932844	C	ADD	5166	0.2843	5.472e-06
1	rs2258497	167408073	C	ADD	5184	2.906	5.619e-06
15	rs17270362	61050486	A	ADD	5187	2.599	6.214e-06
1	rs870873	167399503	G	ADD	5186	2.973	6.805e-06
1	rs953809	167403625	A	ADD	5186	2.995	6.941e-06
4	rs16893097	16156964	A	ADD	5139	0.2208	8.645e-06

15	rs12437690	61055432	A	ADD	5184	2.553	9.529e-06
16	rs11076173	56919235	A	ADD	5180	2.913	1.005e-05
9	rs1403561	15011825	G	ADD	5184	2.718	1.023e-05
9	rs784935	116434346	C	ADD	5139	0.296	1.215e-05
8	rs2006919	17975812	A	ADD	5147	3.528	1.546e-05
3	rs1199333	138091701	C	ADD	5166	0.3308	1.707e-05
7	rs219254	150992226	A	ADD	5180	2.469	1.811e-05

CHR, chromosome; SNP, SNP name; BP, SNP position in base-pairs; A1, Minor Allele; ADD, genetic model additive trend test; OR, Odds Ratios; P, p-value.

5.3.2.2 LocusZoom plots for the two top hits with p<5e-8

As can be seen in our top hits table, the first hit (SNP name: rs6657114) is located on chromosome 1, and the reported p-value in the table is 7.008e-19. Figure 5.5 is a LocusZoom plot (http://locuszoom.sph.umich.edu/locuszoom/) of the above hit, but unfortunately the SNP is not well supported by other SNPs. The LocusZoom plot shows that this is a "singleton" signal, and that SNPs in high linkage disequilibrium with the hit do not show a signal.

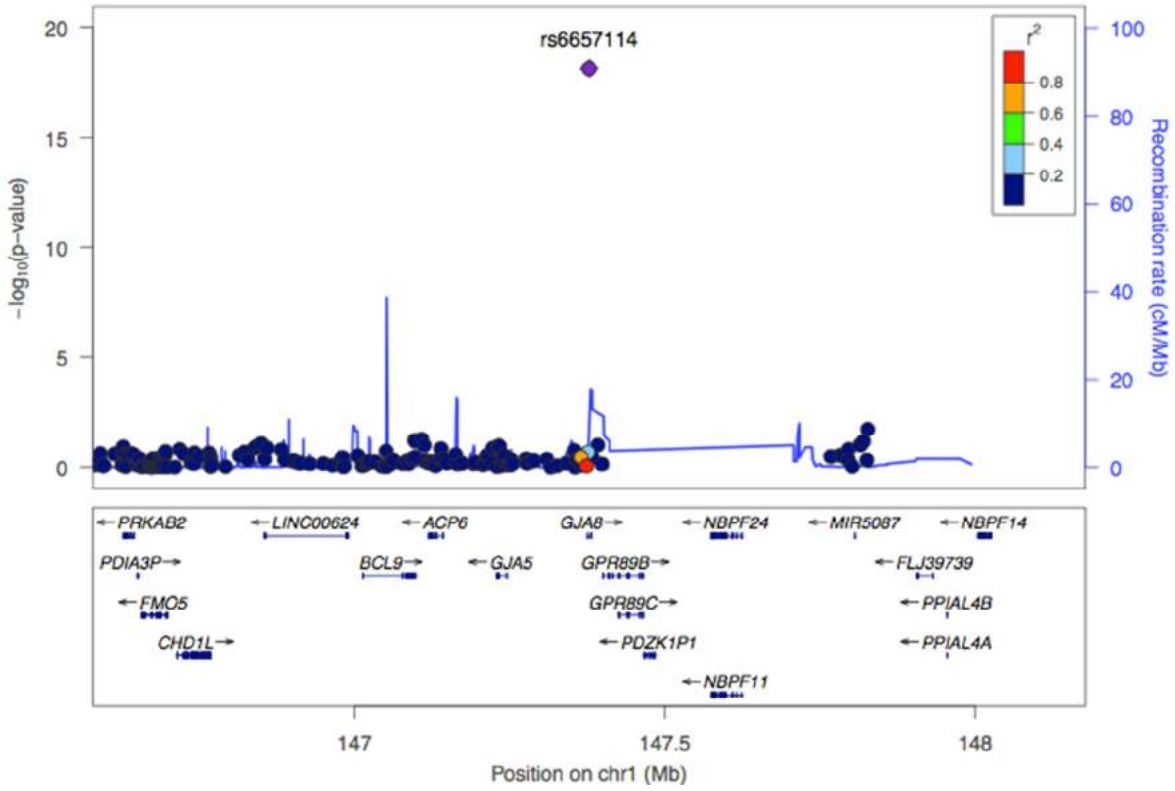


Figure 5.5 LocusZoom plot for the first SNP on the top hits list.
The main panel shows association p-values on the –log10 scale on the vertical axis, and the chromosomal position along the horizontal axis. Genes within the region are shown in the lower panel. The SNP is not supported with other SNPs, therefore the signal is likely not a real hit.

The next hit in our top hits table (SNP name: rs12404733) is also located on chromosome 1, and the reported p-value is 2.684e-08. Figure 5.6 is a LocusZoom plot of the above hit. The LocusZoom plot shows that this is also a "singleton" SNP that is not supported by other SNPs including one that is in high LD.

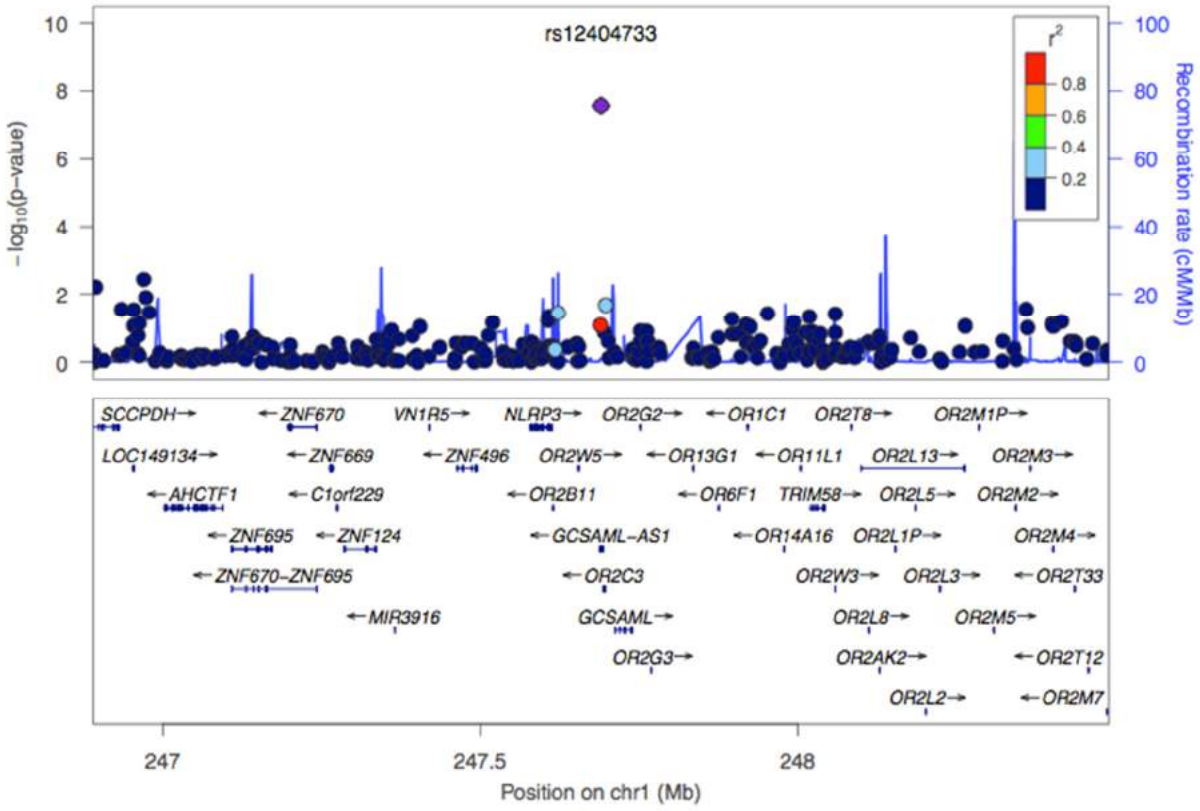


Figure 5.6 LocusZoom plot for the 2nd top hit. The SNP is not supported by other SNPs, therefore the signal is likely not a real hit.
The main panel shows association p-values on the –log10 scale on the vertical axis, and the chromosomal position along the horizontal axis. Genes within the region are shown in the lower panel. Each filled circle represents the p-value for one SNP in the GWAS scan.

5.3.2.3 LocusZoom plots for hits with 5e-8<p<5e-6

I created LocusZoom plots for all hits with p<5e-6, but unfortunately none of the SNPs are well supported by other SNPs and they appear to be "singleton" signals. This pattern is consistent with these hits being false positive signals generated by systematic biases in genotype calling between cases and controls. This

presumably results from the fact that the cases and controls were collected and processed separately and typed on different genotyping platforms.

5.3.2.4 Checking the hit for the TwinsUK GWAS in our clinical GWAS

We checked the hit SNP (rs12438477) from the TwinsUK data for any evidence of significant association, but unfortunately the same SNP is not present in the clinical GWAS result. We therefore checked for any possible LD proxy. By using the SNAP tool (http://www.broadinstitute.org/mpg/snap/), we found out that SNP rs12438477 is on the Illumina 1M chip and there are no other LD proxies ($r2<0.8$) on any Illumina chips. It is likely that the SNP was removed during the QC steps.

Therefore, to look for any evidence of an association signal in the GWAS scan on all cases, we created a LocusZoom plot (Figure 5.7) of the region around rs12438477. As shown in Figure 5.7 there is no evidence of a signal at the same region.

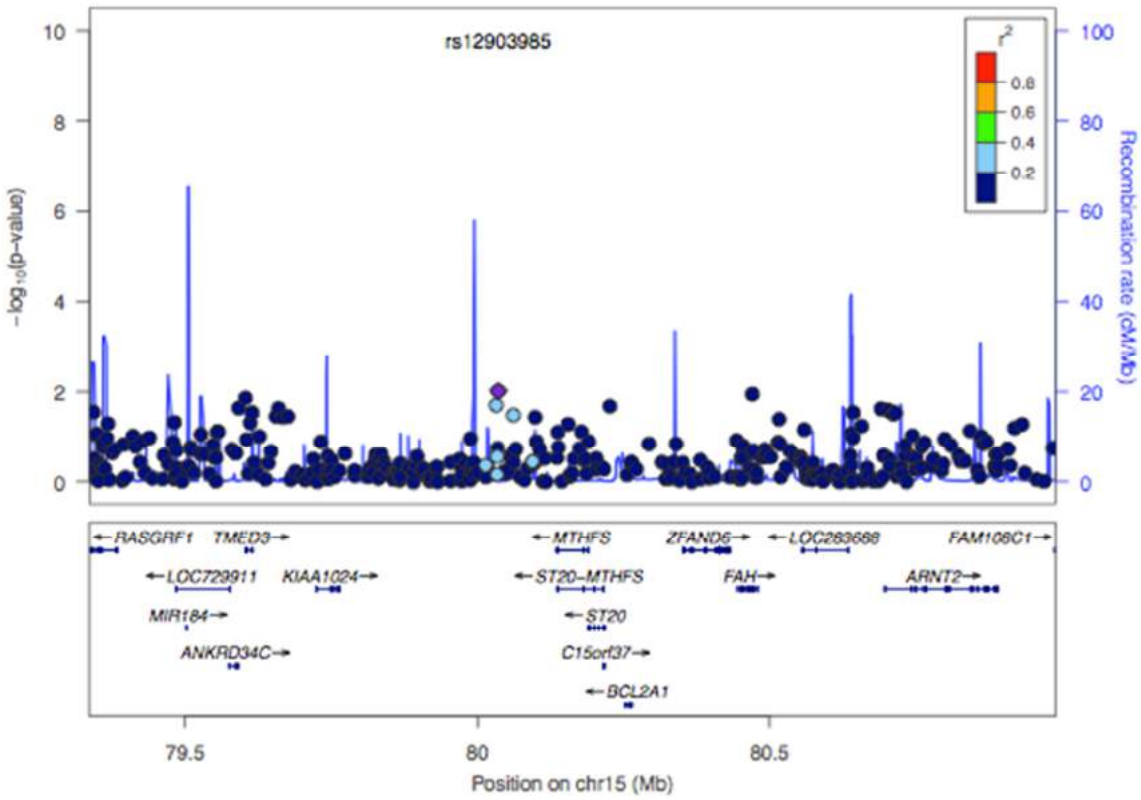


Figure 5.7 LocusZoom plot for the hit from TwinsUK GWAS.
There is no evidence for an association signal in the clinical GWAS data. The main panel shows association p-values on the –log10 scale on the vertical axis, and the chromosomal position along the horizontal axis. Genes within the region are shown in the lower panel. Each filled circle represents the p-value for one SNP in the clinical GWAS scan on all cases.

5.3.3 Checking the signals of other candidate gene studies in our GWAS scans

In order to check if any of the significant SNPs from previous studies are replicated in our data, we created a LocusZoom plot for each of the candidate genes and significant SNPs from previously published studies using our data. Table 5.5 compares the findings of previously published studies with our all-

cases GWAS, amoxicillin-only GWAS and TwinsUK GWAS. Unfortunately the previously reported positive results were not replicated in any of our data sets.

Table 5.5 Comparison of previously published findings with our data

Gene	Variant	Chr	Pos	Rep in Clin GWAS	Rep in Twins GWAS	Rep in Amox GWAS	Notes
HLA-DRA	rs7192	6	32519624	NO	NO	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
C5	rs17612	9	1222765747	NO	NO	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
ZNF300	rs4958427	5	150258780	NO	NO	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs7754768	6	32528157	NO	NO	NO	Discovery hit in 2014 Immunochip study, but not genotyped in replication phase (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs9268832	6	32535767	NO	NO	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs2227139	6	32521437	NO	NO	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs2213586	6	32521072	NO	NO	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA / HLA-DRB5	rs2213585	6	32521128	NO	NO	NO	Candidate in candidate gene study (Gueant et al., 2015)

HLA-DRA	Rs7195	6	32520517	NO	NO	NO	Candidate in candidate gene study (Gueant et al., 2015)
HLA-DRA	rs8084	6	32519013	NO	NO	NO	Replicated in IP in 2014 Immunochip study (Gueant et al., 2015)
NOD2	rs2066845	-	-	NO	NO	NO	Candidate in candidate gene study (Oussalah et al., 2016)
NOD1	rs2907749	-	-	NO	NS	NO	Candidate in candidate gene study (Oussalah et al., 2016)
IL4	Candidate gene	5	-	NO	NO	NO	(Gueant-Rodriguez et al (2006)
IL4R	Candidate gene	-	-	NO	NO	NO	(Gueant-Rodriguez et al., 2006)
IL10	Candidate gene	-	-	NO	NO	NO	(Guglielmi et al., 2006)
IL13	Candidate gene	5	-	NO	NO	NO	(Gueant-Rodriguez et al., 2006)
LACTB	Candidate gene	-	-	NO	NO	NO	(Apter et al., 2008)
TNFA	Candidate gene	-	-	NO	NO	NO	(Cornejo-Garcia et al., 2012)
IL18	Candidate gene	-	-	NO	NO	NO	(Ming et al., 2011)

Table showing the comparison of previously published findings with our data. All the significant SNPs and candidate genes in previous beta-lactam allergy studies have been checked for replication in all three data set (all-cases scan, amoxicillin-only scan, TwinsUK scan). Unfortunately none of the previously reported SNPs or genes are replicated in any of our data sets.

5.3.4 Results for amoxicillin-only scan

5.3.4.1 Top hits for GWAS scan on amoxicillin-only cases

Table 5.6 shows the first 30 top hits for the amoxicillin-only scan. Unfortunately, none of the SNPs achieve genome-wide significance.

Table 5.6 The top 30 hits for the GWAS scan of amoxicillin-only scan.

CHR	SNP	BP	A1	F_A	F_U	A2	P	OR
1	rs2840945	225859141	C	0	0.4616	A	2.27E-06	0
20	rs4925203	60451598	A	0	0.4415	G	3.19E-06	0
2	rs2121434	149557825	G	0.6818	0.2295	A	8.39E-06	7.193
18	rs4941279	62181984	A	0.8182	0.3668	G	1.96E-05	7.77
11	rs11216667	117876198	A	0	0.4025	C	2.28E-05	0
1	rs10798986	34241397	C	0.5455	0.1558	A	2.91E-05	6.502
1	rs6425668	181119835	G	0	0.3817	A	3.02E-05	0
17	rs7216522	43108705	G	0.8636	0.4234	A	3.25E-05	8.625
14	rs12894797	80792090	G	0.9091	0.4919	A	5.30E-05	10.33
14	rs12896139	80792909	G	0.9091	0.4939	A	5.60E-05	10.25
2	rs6739054	11220562	G	0.6818	0.2672	A	5.91E-05	5.877
6	rs2816369	53088181	A	0.4091	0.0896	C	6.46E-05	7.034
15	rs10902565	100773585	A	0.6818	0.2699	G	6.72E-05	5.796
15	rs16954285	70499098	A	0.4545	0.1151	G	7.30E-05	6.404
1	rs6657114	147377718	A	0.2727	0.03403	C	7.59E-05	10.65
14	rs12050217	96728753	G	0.5909	0.2026	A	7.65E-05	5.685
15	rs10459644	70512623	A	0.4091	0.09373	G	9.18E-05	6.694
16	rs11862936	5961636	G	0.7273	0.3162	A	9.21E-05	5.768
3	rs4679561	59422456	A	0.8182	0.4029	G	9.69E-05	6.67
2	rs7580383	125457240	G	0.04545	0.4387	A	0.0001025	0.06093
1	rs621070	218337199	A	0.3636	0.07278	G	0.0001027	7.28
11	rs7937026	96634323	A	0.2273	0.02209	G	0.0001071	13.02
2	rs4499362	149568396	A	0.5909	0.2118	G	0.0001238	5.376
17	rs916660	43158674	A	0.8636	0.4565	G	0.0001245	7.54
18	rs8092218	62191653	G	0.8636	0.4598	A	0.0001336	7.44
1	rs12048806	192410674	A	0.4091	0.09866	G	0.0001365	6.325
2	rs12692166	236258449	A	0	0.3533	G	0.0001386	0
5	rs6883772	158131574	G	0	0.3423	A	0.000144	0
7	rs1155597	126613163	A	0.7727	0.359	C	0.0001474	6.07
18	rs2868934	10204383	A	0.5	0.153	G	0.000149	5.537

CHR=chromosome; SNP=SNP rsID; BP=position (Build37); A1=Allele1; F_A(F_U)=case(control) allele freq; Allele 2; P=p-value; OR=Odds ratio

5.3.4.2 QQ plot for the amoxicillin-only GWAS scan.

Figure 5.8 is the QQ plot for the amoxicillin-only scan. All the SNPs in the upper end are still in the grey zone area (95% concentration band), indicating a lack of any genome-wide significant hit in the results. The QQ plot in figure 5.8 is slightly deflated, which is typical of FET tests. Overall, there is little evidence for any real GWAS hits here, reflecting the fundamental problem with lack of power with attempting to do a GWAS with only 11 cases.

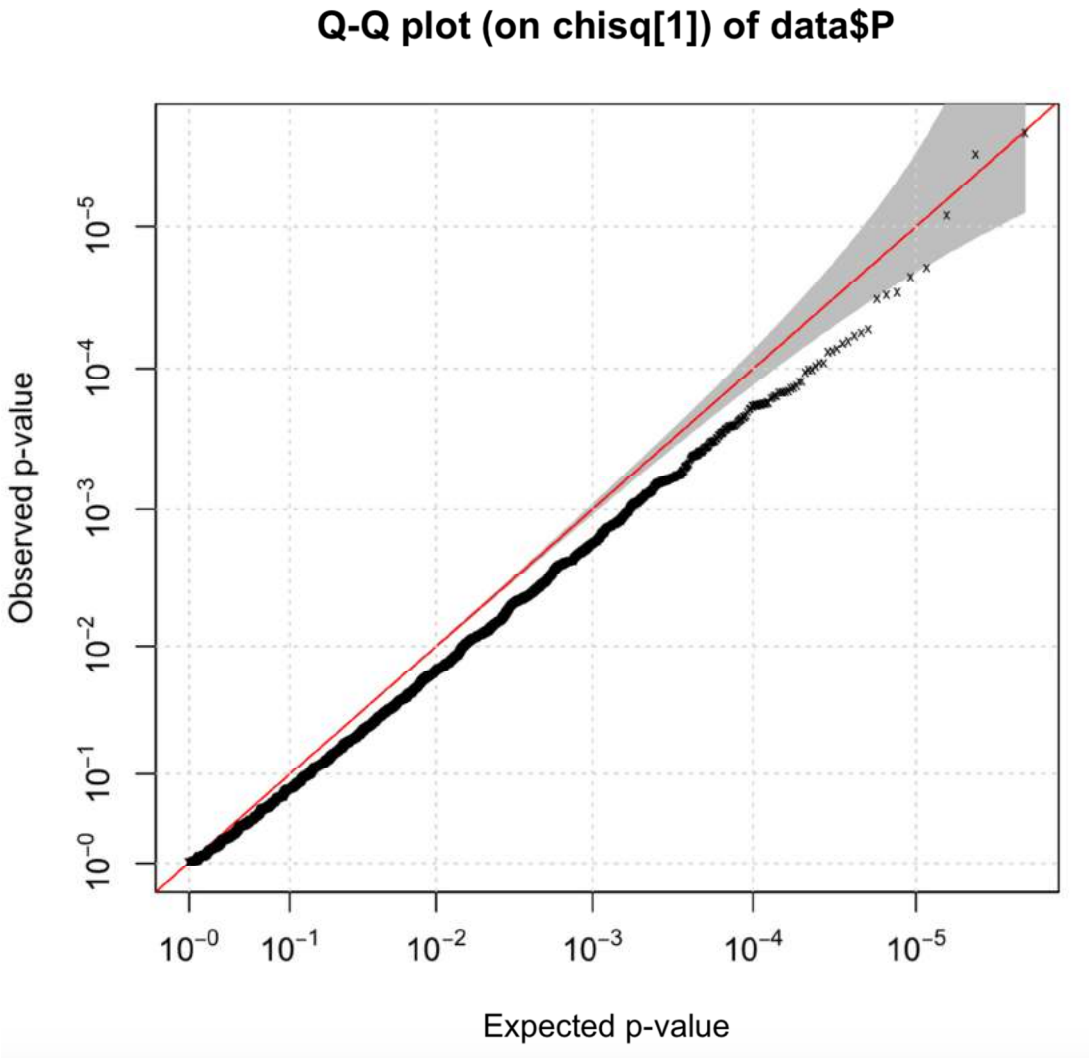


Figure 5.8 QQ plot for the amoxicillin-only scan.
The negative logarithms of the observed (y axis) and the expected (x axis) ordered p-values are plotted for each SNP. The red line indicates the expected distribution under the global null hypothesis. The grey shaded region represents the 95% concentration band. There is no genomic inflation in lower end of the plot. There is no evidence of any significant hits.

5.3.4.3 Manhattan plot for the amoxicillin-only GWAS scan

Figure 5.9 displays the Manhattan plot for the amoxicillin-only GWAS scan. There are no SNPs passing the genome-wide significant threshold, again showing that there are no significant hits in this scan.

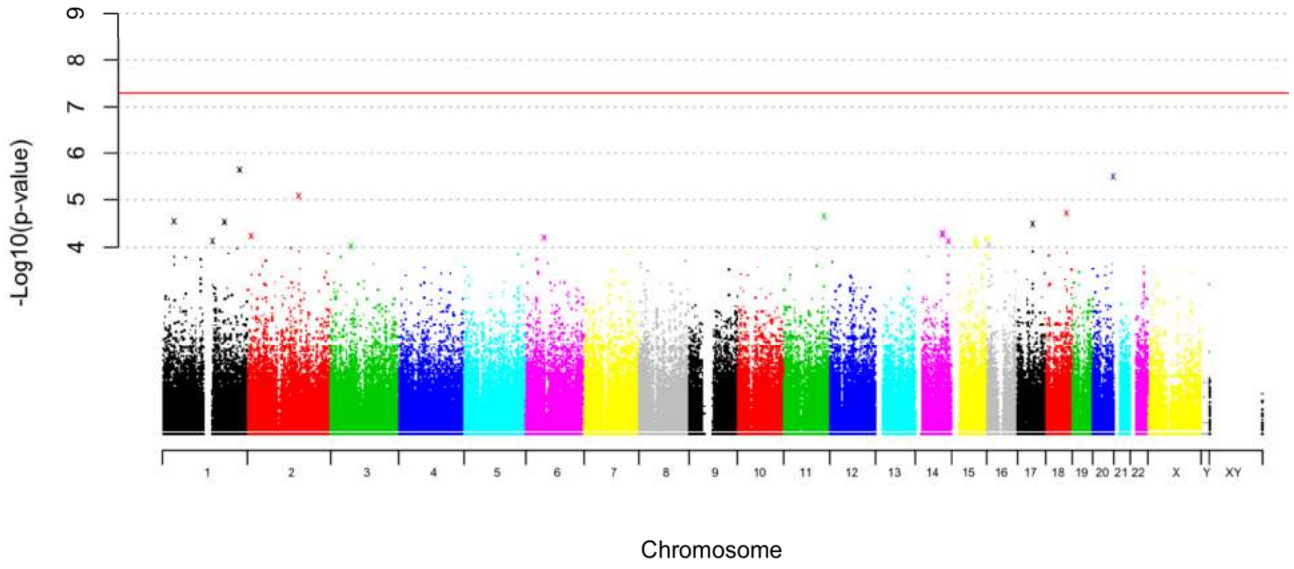


Figure 5.9 Manhattan plot for the amoxicillin-only GWAS scan.
The x-axis shows the chromosomal position and the y-axis shows the –log10 p-value. Chromosomes are shown in alternate colours. The red horizontal line indicates a threshold of genome-wide significance at a p-value of p<5e-8. Each point represents a p-value for the SNP association test. There is no evidence for significant hits.

5.4 Discussion

5.4.1 GWAS scans

Cases recruited from Guy's Hospital allergy clinic are similar to the previous cohort study used by Gueant et al., 2015. In both studies allergic work-up was performed according to the European Network for Drug allergy guidelines. Patients have been recruited after confirming of having a positive history of immediate beta-lactam allergy. Skin tests were performed with major and minor determinants of benzyl penicillin, amoxicillin and culprit drug when a cephalosporin involved. In the case of negative skin test, provocation test was performed with culprit drug in order to complete diagnosis. Also all recruited patients were white.

We have conducted one of the first GWAS studies of well-defined beta-lactam allergy. There is another ongoing similar unpublished study lead by Dr. Munir Pirmohamed at Liverpool University. Despite the lack of well-defined whole genome studies on beta-lactam allergies, the positive findings from the TwinsUK GWAS, which I have previously discussed in Chapter 3, encouraged me to run another GWAS by using clinically well-defined cases in the hope that we could replicate the positive result of the TwinsUK GWAS.

I therefore recruited allergic cases from the Guy's allergy clinic. This study was conducted with 48 cases (see methods section). We used the large WTCCC2 dataset as our control group (see methods section). We included 48 cases, which is comparatively low but, as I discussed in the Introduction chapter, there are examples of successful pharmacogenetics GWAS studies which also had a low number of cases.

Our case recruitment protocol is similar to that of Gueant et al. (2015). In both studies, allergic work-up was performed according to the European Network for Drug Allergy guidelines. Patients were recruited after confirming a positive history of immediate beta-lactam allergy. Skin tests were performed with major and minor determinants of benzyl penicillin, amoxicillin and the culprit drug when a cephalosporin was involved. In the case of a negative skin test, a provocation test was performed with the culprit drug in order to complete diagnosis. Also, all recruited patients were white.

In light of previous successful GWAS studies in the pharmacogenomic field with limited case numbers, it was hoped that by using well-defined cases and a large control group we would potentially have a study that was sufficiently powered to detect a strong pharmacogenetic signal. However, no evidence of significant hit was obtained either from the combined beta-lactam allergy cases from the amoxicillin allergy cases in their own.

5.4.2 Limitations of the clinical GWAS scan

Here I discuss some of the limitations of this GWAS scan, which may have contributed to the negative results that we obtained.

- Small case sample size;

  As explained before, sample size, especially the number of the smaller sample group (here cases), plays a large role in the power of a study. This was confirmed by our power calculations, which showed that we only had adequate power to detect very large genetic association effects (OR>4.3).

- Control selection;

  In our GWAS we used the WTCCC2 control dataset as our control group. There is no information about allergy status for the WTCCC2 cohort. In the case of using any unselected control data, there is always a possibility of having a number of cases among the control group; and this would especially be a problem for high-prevalence phenotypes, this is a minor problem if the prevalence of cases in the general population is small, as illustrated by previous examples of the successful use of unselected controls for rare adverse drug reaction GWASs (e.g. (Nelson et al., 2009)). Previous studies have suggested that the prevalence of beta-lactam allergy in patients who taking beta-lactam antibiotics is relatively small.

- Batch effects;

  In this study, cases and controls were collected separately and genotyped in different labs and different platforms. Batch effects could be therefore another reason for not having a positive result as batch effects can give rise to false negative as well as false positive results.

- Lack of imputation;

  Due to lack of time we didn't impute our data, however this was a negative study with ~500k SNPs. It is unlikely that we missed any possible signals due to lack of imputation. Imputation is valuable as a fine-mapping tool, but this only applies in cases where an actual signal is detected.

5.4.3 Comparing the TwinsUK GWAS to the clinical GWAS

There were no positive findings in our clinical GWAS, as we could not replicate the results for the self-reported TwinsUK GWAS and we did not generate any new genome-wide significant results after QC. The discrepancies in the findings might be explained by the following factors.

- Sample size;

  There was a small case sample size in the clinical GWAS as compared to TwinsUK GWAS. As explained before, sample size plays a significant role

in the power of a study. In the TwinsUK GWAS we had self-reported phenotypes, but the number of cases was larger than the clinical GWAS.

- Phenotype definition;

  The TwinsUK GWAS was conducted on self-reported beta-lactam allergic cases. However in the clinical GWAS we recruited well defined and clinically tested cases.

- Control selection;

  As explained in chapter 3, for the TwinsUK GWAS, we selected super controls (all controls were selected as being allergy free), but in our clinical GWAS we used the WTCCC2 control dataset. In the case of the WTCCC2 cohort, we did not have any information about allergy status. Therefore there is a possibility of having a number of cases within the control group.

- Mismatching phenotype definition;

  In one of the TwinsUK GWAS scans, we included all delayed cases as well as immediate allergic cases, thus assuming that there are some common pathways in immediate and delayed allergic reactions. However in the clinical GWAS, only immediate allergic cases were included.

- Batch effects;

  In the clinical GWAS, cases and controls were collected separately and genotyped in different labs and different platforms. However data for TwinsUK GWAS were collected and genotyped at the same lab and corrected for any existing batch effects.

5.4.4 Checking the results of the previous studies in the clinical GWAS

As I explained before, many factors, such as small sample size, batch effects and a poorly defined control group, are limitations of this negative study. Taking into account the small case sample size in our study and the lack of power, as a final step, we decided to check all the positive findings of the previously published studies to see if we could replicate those results even if there was only a weaker signal in our study. However as shown in Table5.5, previously published results are not replicated in our clinical GWAS.

To conclude, there were no positive findings in our clinical GWAS, as we could not replicate the results for the TwinsUK GWAS nor any previously published positive studies, and we did not generate any new genome-wide significant results after QC. As more studies are performed on this phenotype, we hope that our data can be incorporated into meta-analyses of different cohorts.

Chapter 6: Discussion

6.1 Strengths and limitations of the study

The investigations presented in this thesis have several strengths and advantages. Having access to the data of a large number of twins in the TwinsUK cohort, and data from the Guy's allergy clinic, was one of the obvious strengths of this study. Twin studies provide us with one of the most powerful designs for heritability estimation. Participants from the allergy clinic have been tested and screened for beta-lactam allergy, out of which 48 were considered for the purpose of this thesis.

Designing and validating a detailed and unique questionnaire for the assessment and phenotyping of beta-lactam allergy was another strong point of this study. This unique questionnaire was used to eliminate an appreciable percentage of the self-reported cases from the allergic category, and proof of the benefit of doing this was demonstrated by the improvement in the heritability estimate that followed from applying the revised phenotype definitions to the TwinsUK cohort.

Our investigations also have some limitations that have been mainly discussed in each chapter. By using the new specific beta-lactam allergy questionnaire we were able to overcome some of the limitations of the old questionnaire. However, a number of limitations still remained with the use of the new questionnaire. These limitations have been discussed in detail in Chapter 2.

The strengths and limitations of the TwinsUK GWAS scan on seriousness as a quantitative trait have been discussed in detail in Chapter 3. The strength and

limitations of MWAS on TwinsUK beta-lactam allergy data have been discussed in detail in Chapter 4. The limitations of the clinical GWAS scan have been discussed in detail in chapter 5.

The main limitation pointed out by several practitioners and researchers during this study was using the self-reported twins cohort, rather than a clinically defined phenotype, a problem that still remained even though I took efforts to improve the quality of information obtained from the second questionnaire I designed.

The other limitation of this study, with respect to the low prevalence of the true allergic cases, was the low number of the confirmed allergic cases in the clinical GWAS, which likely resulted in low GWAS power. Using the WTCCC2 cohort as our control group for the clinical GWAS was another possible imperfection of this study. By using a population-based cohort as our control group, there was a potential of having 'allergic controls' in our study. However, since the prevalence of true beta-lactam allergy is not high, the low case number was probably the more problematic issue.

6.2 Future work

Perhaps the most important research question that can be considered as an extension to the works of this thesis concerns the role and connection of the

genomewide-significant associated signal located in the TwinsUK GWAS, which requires confirmation via a replication study (see discussion chapter 3).

Also of interest would be a follow up study to the MWAS. This would help to establish whether the result reported here is biologically valid or due to unknown batch effects. If the former, it would be of interest to follow up the two unknown metabolite hits in the TwinsUK MWAS, to establish their origin.

The observations in this thesis lead us to accept the hypothesis of beta-lactam allergy as being a complex trait that is under both genetic and environmental factors. In this study, we tried to uncover the genetic component of beta-lactam allergy. As explained before, we demonstrated evidence for the involvement of genetic factors in this trait, but the role of a wide variety of environmental factors was not studied. A detailed epidemiological study to understand the contribution of lifestyle could therefore be another follow-up for this work.

As one of the limitations of this study, I explained how well-defined phenotypes can improve the power of GWAS, and how not having a well-defined control group can affect the results. Therefore future work to find new genetic risk factors could be more focused on the different sub-types of beta-lactam allergy and both the common and distinct biological pathways for them.

The work done in this thesis may also be applied in different ways. For heritability estimation, regardless of the current allergy status, administering

the same specially designed questionnaire for all twins in the TwinsUK cohort will improve still further the quality of the allergy severity phenotype.

Referring the self-reported allergic twins to the allergy clinic and applying the clinical tests (skin and challenge test), or checking the blood samples for specific IgE could also add more accuracy to our result. In this case, re-doing the TwinsUK GWAS by using the truly allergic cases and comparing the result with the self reported GWAS could be of interest.

By applying clinical tests on self-reported twins and having clinically proved cases, we could join our cases from clinical GWAS with clinically proved TwinsUK GWAS cases. Having a high number of cases and using our super-controls from the TwinsUK GWAS, we could run another GWAS with more power. In the next step we could apply metabolomic profiling on our clinical cases and join them with twins' data and apply another MWAS with more power. Future studies are generally required to validate the existence of systematic metabolomic effects.

The findings of the investigations carried out in this thesis may serve as a basis for future research in related topics. The results of the TwinsUK GWAS and MWAS of the thesis, should they be validated in independent populations, could be used as potential biomarkers for safer clinical practice. Another avenue is collaboration with other studies using the data available in this field. Meta-analysis of our data with future studies could be used to overcome the power

issue in this research and facilitate further advances in the genetic epidemiology of allergy to beta-lactam antibiotics.

References:

ABDULNOUR-NAKHOUL, S. M., AL-TAWIL, Y., GYFTOPOULOS, A. A., BROWN, K. L., HANSEN, M., BUTCHER, K. F., EIDELWEIN, A. P., NOEL, R. A., RABON, E., POSTA, A. & NAKHOUL, N. L. 2013. Alterations in junctional proteins, inflammatory mediators and extracellular matrix molecules in eosinophilic esophagitis. *Clin Immunol,* 148**,** 265-278.

ALTMAN, R. B., KROEMER, H. K., MCCARTY, C. A., RATAIN, M. J. & RODEN, D. 2011. Pharmacogenomics: will the promise be fulfilled? *Nat Rev Genet,* 12**,** 69-73.

APTER, A. J., SCHELLEMAN, H., WALKER, A., ADDYA, K. & REBBECK, T. 2008. Clinical and genetic risk factors of self-reported penicillin allergy. *J Allergy Clin Immunol,* 122**,** 152-8.

ARIZA, A., MAYORGA, C., FERNANDEZ, T. D., BARBERO, N., MARTIN-SERRANO, A., PEREZ-SALA, D., SANCHEZ-GOMEZ, F. J., BLANCA, M., TORRES, M. J. & MONTANEZ, M. I. 2015. Hypersensitivity reactions to beta-lactams: relevance of hapten-protein conjugates. *J Investig Allergol Clin Immunol,* 25**,** 12-25.

BALDO, B. A. & TOVEY, E. R. 1989. *Protein blotting : methodology, research and diagnostic applications,* Basel, Karger.

BERTUCCI, C., BARSOTTI, M. C., RAFFAELLI, A. & SALVADORI, P. 2001. Binding properties of human albumin modified by covalent binding of penicillin. *Biochim Biophys Acta,* 1544**,** 386-92.

BLANCA, M., ROMANO, A., TORRES, M. J., FERNANDEZ, J., MAYORGA, C., RODRIGUEZ, J., DEMOLY, P., BOUSQUET, P. J., MERK, H. F., SANZ, M. L., OTT, H. & ATANASKOVIC-MARKOVIC, M. 2009. Update on the evaluation of hypersensitivity reactions to betalactams. *Allergy,* 64**,** 183-93.

BOOMSMA, D., BUSJAHN, A. & PELTONEN, L. 2002. Classical twin studies and beyond. *Nat Rev Genet,* 3**,** 872-82.

BROWN, S. G. 2004. Clinical features and severity grading of anaphylaxis. *J Allergy Clin Immunol,* 114**,** 371-6.

BURSZTEJN, A. C., ROMANO, A., GUEANT-RODRIGUEZ, R. M., CORNEJO, J. A., OUSSALAH, A., CHERY, C., GASTIN, I., GAETA, F., BARBAUD, A., BLANCA, M. & GUEANT, J. L. 2013. Allergy to betalactams and nucleotide-binding oligomerization domain (NOD) gene polymorphisms. *Allergy,* 68**,** 1076-80.

BUSH, W. S. & MOORE, J. H. 2012. Chapter 11: Genome-wide association studies. *PLoS Comput Biol,* 8**,** e1002822.

COOKSON, W. 1999. The alliance of genes and environment in asthma and allergy. *Nature,* 402**,** B5-11.

COOPER, G. M., JOHNSON, J. A., LANGAEE, T. Y., FENG, H., STANAWAY, I. B., SCHWARZ, U. I., RITCHIE, M. D., STEIN, C. M., RODEN, D. M., SMITH, J. D., VEENSTRA, D. L., RETTIE, A. E. & RIEDER, M. J. 2008. A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood,* 112**,** 1022-7.

CORNEJO-GARCIA, J. A., GUEANT-RODRIGUEZ, R. M., TORRES, M. J., BLANCA-LOPEZ, N., TRAMOY, D., ROMANO, A., BLANCA, M. & GUEANT, J. L. 2012. Biological and genetic determinants of atopy are predictors of immediate-type allergy to betalactams, in Spain. *Allergy,* 67**,** 1181-5.

DALY, A. K. 2010a. Genome-wide association studies in pharmacogenomics. *Nat Rev Genet,* 11**,** 241-6.
DALY, A. K. 2010b. Pharmacogenetics and human genetic polymorphisms. *Biochem J,* 429**,** 435-49.
DALY, A. K. 2013. Pharmacogenomics of adverse drug reactions. *Genome Med,* 5**,** 5.
DALY, A. K., DONALDSON, P. T., BHATNAGAR, P., SHEN, Y., PE'ER, I., FLORATOS, A., DALY, M. J., GOLDSTEIN, D. B., JOHN, S., NELSON, M. R., GRAHAM, J., PARK, B. K., DILLON, J. F., BERNAL, W., CORDELL, H. J., PIRMOHAMED, M., AITHAL, G. P., DAY, C. P., STUDY, D. & INTERNATIONAL, S. A. E. C. 2009. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nat Genet,* 41**,** 816-9.
DE SWERT, L. F. 1999. Risk factors for allergy. *Eur J Pediatr,* 158**,** 89-94.
DEMOLY, P. & HILLAIRE-BUYS, D. 2004. Classification and epidemiology of hypersensitivity drug reactions. *Immunol Allergy Clin North Am,* 24**,** 345-56, v.
DEMOLY, P. & ROMANO, A. 2005. Update on beta-lactam allergy diagnosis. *Curr Allergy Asthma Rep,* 5**,** 9-14.
DEMOLY, P., ROMANO, A., BOTELHO, C., BOUSQUET-ROUANET, L., GAETA, F., SILVA, R., RUMI, G., RODRIGUES CERNADAS, J. & BOUSQUET, P. J. 2010. Determining the negative predictive value of provocation tests with beta-lactams. *Allergy,* 65**,** 327-32.
DESCOTES, J. & CHOQUET-KASTYLEVSKY, G. 2001a. Gell and Coombs's classification: is it still valid? *Toxicology,* 158**,** 43-9.
DESCOTES, J. & CHOQUET-KASTYLEVSKY, G. 2001b. Gell and Coombs's classification: is it still valid? *Toxicology,* 158**,** 43-9.
DESHAZO, R. D. & KEMP, S. F. 1997. Allergic reactions to drugs and biologic agents. *JAMA,* 278**,** 1895-906.
DUAN, S., ZHANG, W., COX, N. J. & DOLAN, M. E. 2008. FstSNP-HapMap3: a database of SNPs with high population differentiation for HapMap3. *Bioinformation,* 3**,** 139-41.
EDFORS-LUBS, M. L. 1971. Allergy in 7000 twin pairs. *Acta Allergol,* 26**,** 249-85.
EHRET, G. B. 2010. Genome-wide association studies: contribution of genomics to understanding blood pressure and essential hypertension. *Curr Hypertens Rep,* 12**,** 17-25.
EU-AHSUNTHORNWATTANA, J., MILLER, E. N., FAKIOLA, M., WELLCOME TRUST CASE CONTROL, C., JERONIMO, S. M., BLACKWELL, J. M. & CORDELL, H. J. 2014. Comparison of methods to account for relatedness in genome-wide association studies with family-based data. *PLoS Genet,* 10**,** e1004445.
FITZHARRIS, P. 2008. Penicillin allergy: updating the role of skin testing in diagnosis. *Postgrad Med J,* 84**,** 505-6.
GAO, N., QIAO, H. L., JIA, L. J., TIAN, X. & ZHANG, Y. W. 2008. Relationships between specific serum IgE, IgG, IFN-gamma level and IFN-gamma, IFNR1 polymorphisms in patients with penicillin allergy. *Eur J Clin Pharmacol,* 64**,** 971-7.
GELL PGH, C. R., EDS 1963. Clinical Aspects of Immunology.
GENETIC ANALYSIS OF PSORIASIS, C., THE WELLCOME TRUST CASE CONTROL, C., STRANGE, A., CAPON, F., SPENCER, C. C., KNIGHT, J., WEALE, M. E., ALLEN, M. H., BARTON, A., BAND, G., BELLENGUEZ, C., BERGBOER, J. G., BLACKWELL, J. M., BRAMON, E., BUMPSTEAD, S. J., CASAS, J. P., CORK, M. J., CORVIN, A., DELOUKAS, P., DILTHEY, A., DUNCANSON, A., EDKINS, S., ESTIVILL, X., FITZGERALD, O., FREEMAN, C., GIARDINA, E., GRAY, E., HOFER, A., HUFFMEIER, U., HUNT, S. E., IRVINE, A. D., JANKOWSKI, J., KIRBY, B., LANGFORD, C., LASCORZ, J., LEMAN, J., LESLIE, S., MALLBRIS, L., MARKUS, H. S., MATHEW, C. G., MCLEAN, W. H.,

MCMANUS, R., MOSSNER, R., MOUTSIANAS, L., NALUAI, A. T., NESTLE, F. O., NOVELLI, G., ONOUFRIADIS, A., PALMER, C. N., PERRICONE, C., PIRINEN, M., PLOMIN, R., POTTER, S. C., PUJOL, R. M., RAUTANEN, A., RIVEIRA-MUNOZ, E., RYAN, A. W., SALMHOFER, W., SAMUELSSON, L., SAWCER, S. J., SCHALKWIJK, J., SMITH, C. H., STAHLE, M., SU, Z., TAZI-AHNINI, R., TRAUPE, H., VISWANATHAN, A. C., WARREN, R. B., WEGER, W., WOLK, K., WOOD, N., WORTHINGTON, J., YOUNG, H. S., ZEEUWEN, P. L., HAYDAY, A., BURDEN, A. D., GRIFFITHS, C. E., KERE, J., REIS, A., MCVEAN, G., EVANS, D. M., BROWN, M. A., BARKER, J. N., PELTONEN, L., DONNELLY, P. & TREMBATH, R. C. 2010. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet,* 42**,** 985-90.

GROUP, S. C., LINK, E., PARISH, S., ARMITAGE, J., BOWMAN, L., HEATH, S., MATSUDA, F., GUT, I., LATHROP, M. & COLLINS, R. 2008. SLCO1B1 variants and statin-induced myopathy--a genomewide study. *N Engl J Med,* 359**,** 789-99.

GRUCHALLA, R. S. 2001. Drug metabolism, danger signals, and drug-induced hypersensitivity. *J Allergy Clin Immunol,* 108**,** 475-88.

GUEANT, J. L., ROMANO, A., CORNEJO-GARCIA, J. A., OUSSALAH, A., CHERY, C., BLANCA-LOPEZ, N., GUEANT-RODRIGUEZ, R. M., GAETA, F., ROUYER, P., JOSSE, T., CANTO, G., CARMONA, F. D., BOSSINI-CASTILLO, L., MARTIN, J., LAGUNA, J. J., FERNANDEZ, J., FEO, F., OSTROV, D. A., PLASENCIA, P. C., MAYORGA, C., TORRES, M. J. & BLANCA, M. 2014. HLA-DRA variants predict penicillin allergy in genome-wide fine-mapping genotyping. *J Allergy Clin Immunol*.

GUEANT, J. L., ROMANO, A., CORNEJO-GARCIA, J. A., OUSSALAH, A., CHERY, C., BLANCA-LOPEZ, N., GUEANT-RODRIGUEZ, R. M., GAETA, F., ROUYER, P., JOSSE, T., CANTO, G., CARMONA, F. D., BOSSINI-CASTILLO, L., MARTIN, J., LAGUNA, J. J., FERNANDEZ, J., FEO, F., OSTROV, D. A., PLASENCIA, P. C., MAYORGA, C., TORRES, M. J. & BLANCA, M. 2015. HLA-DRA variants predict penicillin allergy in genome-wide fine-mapping genotyping. *J Allergy Clin Immunol,* 135**,** 253-9.

GUEANT-RODRIGUEZ, R. M., GUEANT, J. L., VIOLA, M., TRAMOY, D., GAETA, F. & ROMANO, A. 2008. Association of tumor necrosis factor-alpha -308G>A polymorphism with IgE-mediated allergy to betalactams in an Italian population. *Pharmacogenomics J,* 8**,** 162-8.

GUEANT-RODRIGUEZ, R. M., ROMANO, A., BERI-DEXHEIMER, M., VIOLA, M., GAETA, F. & GUEANT, J. L. 2006. Gene-gene interactions of IL13 and IL4RA variants in immediate allergic reactions to betalactam antibiotics. *Pharmacogenet Genomics,* 16**,** 713-9.

GUGLIELMI, L., FONTAINE, C., GOUGAT, C., AVINENS, O., ELIAOU, J. F., GUGLIELMI, P. & DEMOLY, P. 2006. IL-10 promoter and IL4-Ralpha gene SNPs are associated with immediate beta-lactam allergy in atopic women. *Allergy,* 61**,** 921-7.

HIERSCHE, M., RUHLE, F. & STOLL, M. 2013. Postgwas: advanced GWAS interpretation in R. *PLoS One,* 8**,** e71775.

HILL, W. G. & MACKAY, T. F. 2004. D. S. Falconer and Introduction to quantitative genetics. *Genetics,* 167**,** 1529-36.

HINDS, D. A., MCMAHON, G., KIEFER, A. K., DO, C. B., ERIKSSON, N., EVANS, D. M., ST POURCAIN, B., RING, S. M., MOUNTAIN, J. L., FRANCKE, U., DAVEY-SMITH, G., TIMPSON, N. J. & TUNG, J. Y. 2013. A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci. *Nat Genet,* 45**,** 907-11.

HOLM, A. & MOSBECH, H. 2011. Challenge Test Results in Patients With Suspected Penicillin Allergy, but No Specific IgE. *Allergy Asthma Immunol Res,* 3**,** 118-22.

HONG, X., HAO, K., LADD-ACOSTA, C., HANSEN, K. D., TSAI, H. J., LIU, X., XU, X., THORNTON, T. A., CARUSO, D., KEET, C. A., SUN, Y., WANG, G., LUO, W., KUMAR,

R., FULEIHAN, R., SINGH, A. M., KIM, J. S., STORY, R. E., GUPTA, R. S., GAO, P., CHEN, Z., WALKER, S. O., BARTELL, T. R., BEATY, T. H., FALLIN, M. D., SCHLEIMER, R., HOLT, P. G., NADEAU, K. C., WOOD, R. A., PONGRACIC, J. A., WEEKS, D. E. & WANG, X. 2015. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. *Nat Commun,* 6**,** 6304.

HOWIE, B. N., DONNELLY, P. & MARCHINI, J. 2009. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet,* 5**,** e1000529.

HU, Y., SHMYGELSKA, A., TRAN, D., ERIKSSON, N., TUNG, J. Y. & HINDS, D. A. 2016. GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nat Commun,* 7**,** 10448.

HUANG, C. Z., YANG, J., QIAO, H. L. & JIA, L. J. 2009. Polymorphisms and haplotype analysis of IL-4Ralpha Q576R and I75V in patients with penicillin allergy. *Eur J Clin Pharmacol,* 65**,** 895-902.

HUANG, C. Z., ZOU, D., YANG, J. & QIAO, H. L. 2012. Polymorphisms of STAT6 and specific serum IgE levels in patients with penicillin allergy. *Int J Clin Pharmacol Ther,* 50**,** 461-7.

ILLIG, T., GIEGER, C., ZHAI, G., ROMISCH-MARGL, W., WANG-SATTLER, R., PREHN, C., ALTMAIER, E., KASTENMULLER, G., KATO, B. S., MEWES, H. W., MEITINGER, T., DE ANGELIS, M. H., KRONENBERG, F., SORANZO, N., WICHMANN, H. E., SPECTOR, T. D., ADAMSKI, J. & SUHRE, K. 2010. A genome-wide perspective of genetic variation in human metabolism. *Nat Genet,* 42**,** 137-41.

JAMES, C. W. & GURK-TURNER, C. 2001. Cross-reactivity of beta-lactam antibiotics. *Proc (Bayl Univ Med Cent),* 14**,** 106-7.

JOHANSSON, S. G., HOURIHANE, J. O., BOUSQUET, J., BRUIJNZEEL-KOOMEN, C., DREBORG, S., HAAHTELA, T., KOWALSKI, M. L., MYGIND, N., RING, J., VAN CAUWENBERGE, P., VAN HAGE-HAMSTEN, M., WUTHRICH, B. & FORCE, E. N. T. 2001. A revised nomenclature for allergy. An EAACI position statement from the EAACI nomenclature task force. *Allergy,* 56**,** 813-24.

JOHNSON, J. A. & CAVALLARI, L. H. 2015. Warfarin pharmacogenetics. *Trends Cardiovasc Med,* 25**,** 33-41.

KAISER, H. B. 2004. Risk factors in allergy/asthma. *Allergy Asthma Proc,* 25**,** 7-10.

KASTENMULLER, G., RAFFLER, J., GIEGER, C. & SUHRE, K. 2015. Genetics of human metabolism: an update. *Hum Mol Genet,* 24**,** R93-R101.

KELKAR, P. S. & LI, J. T. 2001. Cephalosporin allergy. *N Engl J Med,* 345**,** 804-9.

KHAN, D. A. & SOLENSKY, R. 2010. Drug allergy. *J Allergy Clin Immunol,* 125**,** S126-37.

KOBAYASHI, E. & SATOH, N. 2009. Public involvement in pharmacogenomics research: a national survey on public attitudes towards pharmacogenomics research and the willingness to donate DNA samples to a DNA bank in Japan. *Cell Tissue Bank,* 10**,** 281-91.

KROEMER, H. K. & MEYER ZU SCHWABEDISSEN, H. E. 2010. A piece in the puzzle of personalised medicine. *Clin Pharmacol Ther,* 87**,** 19-20.

LIN, E., SAXON, A. & RIEDL, M. 2010. Penicillin allergy: value of including amoxicillin as a determinant in penicillin skin testing. *Int Arch Allergy Immunol,* 152**,** 313-8.

LIN, J. J., CHANG, C. H. & PAL, N. 2015. A revisit to contingency table and tests of independence: bootstrap is preferred to Chi-square approximations as well as Fisher's exact test. *J Biopharm Stat,* 25**,** 438-58.

LOPEZ-DE-ANDRES, A., JIMENEZ-TRUJILLO, M. I., HERNANDEZ-BARRERA, V., DE MIGUEL-YANES, J. M., MENDEZ-BAILON, M., PEREZ-FARINOS, N., DE BURGOS LUNAR, C., CARDENAS-VALLADOLID, J., SALINERO-FORT, M. A., JIMENEZ-GARCIA, R. & CARRASCO-GARRIDO, P. 2016. Correction: Trends in the

Prevalence of Depression in Hospitalized Patients with Type 2 Diabetes in Spain: Analysis of Hospital Discharge Data from 2001 to 2011. *PLoS One,* 11**,** e0162227.

LUCENA, M. I., MOLOKHIA, M., SHEN, Y., URBAN, T. J., AITHAL, G. P., ANDRADE, R. J., DAY, C. P., RUIZ-CABELLO, F., DONALDSON, P. T., STEPHENS, C., PIRMOHAMED, M., ROMERO-GOMEZ, M., NAVARRO, J. M., FONTANA, R. J., MILLER, M., GROOME, M., BONDON-GUITTON, E., CONFORTI, A., STRICKER, B. H., CARVAJAL, A., IBANEZ, L., YUE, Q. Y., EICHELBAUM, M., FLORATOS, A., PE'ER, I., DALY, M. J., GOLDSTEIN, D. B., DILLON, J. F., NELSON, M. R., WATKINS, P. B., DALY, A. K., SPANISH, D. R., EUDRAGENE, DILIN, DILIGEN & INTERNATIONAL, S. 2011. Susceptibility to amoxicillin-clavulanate-induced liver injury is influenced by multiple HLA class I and II alleles. *Gastroenterology,* 141**,** 338-47.

MA, Q. & LU, A. Y. 2011. Pharmacogenetics, pharmacogenomics, and individualized medicine. *Pharmacol Rev,* 63**,** 437-59.

MALLAL, S., PHILLIPS, E., CAROSI, G., MOLINA, J. M., WORKMAN, C., TOMAZIC, J., JAGEL-GUEDES, E., RUGINA, S., KOZYREV, O., CID, J. F., HAY, P., NOLAN, D., HUGHES, S., HUGHES, A., RYAN, S., FITCH, N., THORBORN, D., BENBOW, A. & TEAM, P.-S. 2008. HLA-B*5701 screening for hypersensitivity to abacavir. *N Engl J Med,* 358**,** 568-79.

MARTINEZ, J. L., FAJARDO, A., GARMENDIA, L., HERNANDEZ, A., LINARES, J. F., MARTINEZ-SOLANO, L. & SANCHEZ, M. B. 2009. A global view of antibiotic resistance. *FEMS Microbiol Rev,* 33**,** 44-65.

MCCORMACK, M., ALFIREVIC, A., BOURGEOIS, S., FARRELL, J. J., KASPERAVICIUTE, D., CARRINGTON, M., SILLS, G. J., MARSON, T., JIA, X., DE BAKKER, P. I., CHINTHAPALLI, K., MOLOKHIA, M., JOHNSON, M. R., O'CONNOR, G. D., CHAILA, E., ALHUSAINI, S., SHIANNA, K. V., RADTKE, R. A., HEINZEN, E. L., WALLEY, N., PANDOLFO, M., PICHLER, W., PARK, B. K., DEPONDT, C., SISODIYA, S. M., GOLDSTEIN, D. B., DELOUKAS, P., DELANTY, N., CAVALLERI, G. L. & PIRMOHAMED, M. 2011. HLA-A*3101 and carbamazepine-induced hypersensitivity reactions in Europeans. *N Engl J Med,* 364**,** 1134-43.

MESSAAD, D., SAHLA, H., BENAHMED, S., GODARD, P., BOUSQUET, J. & DEMOLY, P. 2004. Drug provocation tests in patients with a history suggesting an immediate drug hypersensitivity reaction. *Ann Intern Med,* 140**,** 1001-6.

MING, L., WEN, Q., QIAO, H. L. & DONG, Z. M. 2011. Interleukin-18 and IL18 -607A/C and -137G/C gene polymorphisms in patients with penicillin allergy. *J Int Med Res,* 39**,** 388-98.

MIRAKIAN, R., LEECH, S. C., KRISHNA, M. T., RICHTER, A. G., HUBER, P. A., FAROOQUE, S., KHAN, N., PIRMOHAMED, M., CLARK, A. T., NASSER, S. M., STANDARDS OF CARE COMMITTEE OF THE BRITISH SOCIETY FOR, A. & CLINICAL, I. 2015. Management of allergy to penicillins and other beta-lactams. *Clin Exp Allergy,* 45**,** 300-27.

MOAYYERI, A., HAMMOND, C. J., HART, D. J. & SPECTOR, T. D. 2013. The UK Adult Twin Registry (TwinsUK Resource). *Twin Res Hum Genet,* 16**,** 144-9.

MORRIS, D. E., PEPE, M. S. & BARLOW, W. E. 2010. Contrasting two frameworks for ROC analysis of ordinal ratings. *Med Decis Making,* 30**,** 484-98.

NAG, A., VENTURINI, C., SMALL, K. S., INTERNATIONAL GLAUCOMA GENETICS, C., YOUNG, T. L., VISWANATHAN, A. C., MACKEY, D. A., HYSI, P. G. & HAMMOND, C. 2014. A genome-wide association study of intra-ocular pressure suggests a novel association in the gene FAM125B in the TwinsUK cohort. *Hum Mol Genet,* 23**,** 3343-8.

NELSON, M. R., BACANU, S. A., MOSTELLER, M., LI, L., BOWMAN, C. E., ROSES, A. D., LAI, E. H. & EHM, M. G. 2009. Genome-wide approaches to identify pharmacogenetic contributions to adverse drug reactions. Pharmacogenomics J, 9, 23-33.

NGUYEN, J. K., FOUTS, M. M., KOTABE, S. E. & LO, E. 2006. Polypharmacy as a risk factor for adverse drug reactions in geriatric nursing home residents. *Am J Geriatr Pharmacother,* 4**,** 36-41.

NICHOLSON, J. K., CONNELLY, J., LINDON, J. C. & HOLMES, E. 2002. Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discov,* 1**,** 153-61.

NICHOLSON, J. K. & LINDON, J. C. 2008. Systems biology: Metabonomics. *Nature,* 455**,** 1054-6.

NICHOLSON, J. K. & WILSON, I. D. 2003. Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. *Nat Rev Drug Discov,* 2**,** 668-76.

OUSSALAH, A., MAYORGA, C., BLANCA, M., BARBAUD, A., NAKONECHNA, A., CERNADAS, J., GOTUA, M., BROCKOW, K., CAUBET, J. C., BIRCHER, A., ATANASKOVIC, M., DEMOLY, P., L, K. T., TERREEHORST, I., LAGUNA, J. J., ROMANO, A., GUEANT, J. L. & TASK FORCE 'GENETIC PREDICTORS OF DRUG HYPERSENSITIVITY' OF THE EUROPEAN NETWORK ON DRUG ALLERGY OF, E. 2016. Genetic variants associated with drugs-induced immediate hypersensitivity reactions: a PRISMA-compliant systematic review. *Allergy,* 71**,** 443-62.

PARK, B. K., NAISBITT, D. J., GORDON, S. F., KITTERINGHAM, N. R. & PIRMOHAMED, M. 2001. Metabolic activation in drug allergies. *Toxicology,* 158**,** 11-23.

PETERSEN, A. K., ZEILINGER, S., KASTENMULLER, G., ROMISCH-MARGL, W., BRUGGER, M., PETERS, A., MEISINGER, C., STRAUCH, K., HENGSTENBERG, C., PAGEL, P., HUBER, F., MOHNEY, R. P., GRALLERT, H., ILLIG, T., ADAMSKI, J., WALDENBERGER, M., GIEGER, C. & SUHRE, K. 2014. Epigenetics meets metabolomics: an epigenome-wide association study with blood serum metabolic traits. *Hum Mol Genet,* 23**,** 534-45.

PICHLER, W. J. (ed.) 2007a. *Drug Hypersensitivity*.

PICHLER, W. J. (ed.) 2007b. *Drug Hypersensitivity*.

PICHLER, W. J. 2008. The p-i Concept: Pharmacological Interaction of Drugs With Immune Receptors. *World Allergy Organ J,* 1**,** 96-102.

PICHLER, W. J., ADAM, J., DAUBNER, B., GENTINETTA, T., KELLER, M. & YERLY, D. 2010. Drug hypersensitivity reactions: pathomechanism and clinical symptoms. *Med Clin North Am,* 94**,** 645-64, xv.

PICHLER, W. J., NAISBITT, D. J. & PARK, B. K. 2011. Immune pathomechanism of drug hypersensitivity reactions. *J Allergy Clin Immunol,* 127**,** S74-81.

PIRMOHAMED, M. 2006. Genetic factors in the predisposition to drug-induced hypersensitivity reactions. *AAPS J,* 8**,** E20-6.

PIRMOHAMED, M. 2014. Personalised pharmacogenomics: predicting efficacy and adverse drug reactions. *Annu Rev Genomics Hum Genet,* 15**,** 349-70.

PIRMOHAMED, M., BRECKENRIDGE, A. M., KITTERINGHAM, N. R. & PARK, B. K. 1998. Adverse drug reactions. *BMJ,* 316**,** 1295-8.

PIRMOHAMED, M., JAMES, S., MEAKIN, S., GREEN, C., SCOTT, A. K., WALLEY, T. J., FARRAR, K., PARK, B. K. & BRECKENRIDGE, A. M. 2004. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ,* 329**,** 15-9.

PIRMOHAMED, M. & PARK, B. K. 2003. Adverse drug reactions: back to the future. *Br J Clin Pharmacol,* 55**,** 486-92.

PRUIM, R. J., WELCH, R. P., SANNA, S., TESLOVICH, T. M., CHINES, P. S., GLIEDT, T. P., BOEHNKE, M., ABECASIS, G. R. & WILLER, C. J. 2010. LocusZoom: regional

visualization of genome-wide association scan results. *Bioinformatics,* 26**,** 2336-7.
QIAO, H. L., WEN, Q., GAO, N., TIAN, X. & JIA, L. J. 2007. Association of IL-10 level and IL-10 promoter SNPs with specific antibodies in penicillin-allergic patients. *Eur J Clin Pharmacol,* 63**,** 263-9.
QIAO, H. L., YANG, J. & ZHANG, Y. W. 2004. Specific serum IgE levels and FcepsilonRIbeta genetic polymorphism in patients with penicillins allergy. *Allergy,* 59**,** 1326-32.
QIAO, H. L., YANG, J. & ZHANG, Y. W. 2005. Relationships between specific serum IgE, cytokines and polymorphisms in the IL-4, IL-4Ralpha in patients with penicillins allergy. *Allergy,* 60**,** 1053-9.
RAJAN, T. V. 2003a. The Gell-Coombs classification of hypersensitivity reactions: a re-interpretation. *Trends Immunol,* 24**,** 376-9.
RAJAN, T. V. 2003b. The Gell-Coombs classification of hypersensitivity reactions: a re-interpretation. *Trends Immunol,* 24**,** 376-9.
RICHARDS, J. B., RIVADENEIRA, F., INOUYE, M., PASTINEN, T. M., SORANZO, N., WILSON, S. G., ANDREW, T., FALCHI, M., GWILLIAM, R., AHMADI, K. R., VALDES, A. M., ARP, P., WHITTAKER, P., VERLAAN, D. J., JHAMAI, M., KUMANDURI, V., MOORHOUSE, M., VAN MEURS, J. B., HOFMAN, A., POLS, H. A., HART, D., ZHAI, G., KATO, B. S., MULLIN, B. H., ZHANG, F., DELOUKAS, P., UITTERLINDEN, A. G. & SPECTOR, T. D. 2008. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet,* 371**,** 1505-12.
RIJSDIJK, F. V. & SHAM, P. C. 2002. Analytic approaches to twin data using structural equation models. *Brief Bioinform,* 3**,** 119-33.
RING, J., KRAMER, U., SCHAFER, T. & BEHRENDT, H. 2001. Why are allergies increasing? *Curr Opin Immunol,* 13**,** 701-8.
ROSARIO, N. A. & GRUMACH, A. S. 2006. Allergy to beta-lactams in pediatrics: a practical approach. *J Pediatr (Rio J),* 82**,** S181-8.
SASTRE, J., QUIJANO, L. D., NOVALBOS, A., HERNANDEZ, G., CUESTA, J., DE LAS HERAS, M., LLUCH, M. & FERNANDEZ, M. 1996. Clinical cross-reactivity between amoxicillin and cephadroxil in patients allergic to amoxicillin and with good tolerance of penicillin. *Allergy,* 51**,** 383-86.
SATTA, G., HILL, V., LANZMAN, M. & BALAKRISHNAN, I. 2013. beta-lactam allergy: clinical implications and costs. *Clin Mol Allergy,* 11**,** 2.
SCHNYDER, B. & PICHLER, W. J. 2009. Mechanisms of drug-induced allergy. *Mayo Clin Proc,* 84**,** 268-72.
SHEWBRIDGE, C. 2014. *OECD reviews of evaluation and assessment in education : Northern Ireland, United Kingdom,* Paris, OECD.
SHIINA, T., HOSOMICHI, K., INOKO, H. & KULSKI, J. K. 2009. The HLA genomic loci map: expression, interaction, diversity and disease. *J Hum Genet,* 54**,** 15-39.
SHIN, S. Y., FAUMAN, E. B., PETERSEN, A. K., KRUMSIEK, J., SANTOS, R., HUANG, J., ARNOLD, M., ERTE, I., FORGETTA, V., YANG, T. P., WALTER, K., MENNI, C., CHEN, L., VASQUEZ, L., VALDES, A. M., HYDE, C. L., WANG, V., ZIEMEK, D., ROBERTS, P., XI, L., GRUNDBERG, E., MULTIPLE TISSUE HUMAN EXPRESSION RESOURCE, C., WALDENBERGER, M., RICHARDS, J. B., MOHNEY, R. P., MILBURN, M. V., JOHN, S. L., TRIMMER, J., THEIS, F. J., OVERINGTON, J. P., SUHRE, K., BROSNAN, M. J., GIEGER, C., KASTENMULLER, G., SPECTOR, T. D. & SORANZO, N. 2014. An atlas of genetic influences on human blood metabolites. *Nat Genet,* 46**,** 543-50.
SINGH, A., YAMAMOTO, M., KAM, S. H., RUAN, J., GAUVREAU, G. M., O'BYRNE, P. M., FITZGERALD, J. M., SCHELLENBERG, R., BOULET, L. P., WOJEWODKA, G., KANAGARATHAM, C., DE SANCTIS, J. B., RADZIOCH, D. & TEBBUTT, S. J. 2013. Gene-metabolite expression in blood can discriminate allergen-induced isolated early from dual asthmatic responses. *PLoS One,* 8**,** e67907.

SMALL, K. S., HEDMAN, A. K., GRUNDBERG, E., NICA, A. C., THORLEIFSSON, G., KONG, A., THORSTEINDOTTIR, U., SHIN, S. Y., RICHARDS, H. B., CONSORTIUM, G., INVESTIGATORS, M., CONSORTIUM, D., SORANZO, N., AHMADI, K. R., LINDGREN, C. M., STEFANSSON, K., DERMITZAKIS, E. T., DELOUKAS, P., SPECTOR, T. D., MCCARTHY, M. I. & MU, T. C. 2011. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nat Genet,* 43**,** 561-4.

SMITH, B. N., NEWHOUSE, S., SHATUNOV, A., VANCE, C., TOPP, S., JOHNSON, L., MILLER, J., LEE, Y., TROAKES, C., SCOTT, K. M., JONES, A., GRAY, I., WRIGHT, J., HORTOBAGYI, T., AL-SARRAJ, S., ROGELJ, B., POWELL, J., LUPTON, M., LOVESTONE, S., SAPP, P. C., WEBER, M., NESTOR, P. J., SCHELHAAS, H. J., ASBROEK, A. A., SILANI, V., GELLERA, C., TARONI, F., TICOZZI, N., VAN DEN BERG, L., VELDINK, J., VAN DAMME, P., ROBBERECHT, W., SHAW, P. J., KIRBY, J., PALL, H., MORRISON, K. E., MORRIS, A., DE BELLEROCHE, J., VIANNEY DE JONG, J. M., BAAS, F., ANDERSEN, P. M., LANDERS, J., BROWN, R. H., JR., WEALE, M. E., AL-CHALABI, A. & SHAW, C. E. 2013. The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *Eur J Hum Genet,* 21**,** 102-8.

SOLENSKY, R. 2003. Hypersensitivity reactions to beta-lactam antibiotics. *Clin Rev Allergy Immunol,* 24**,** 201-20.

SORANZO, N., RIVADENEIRA, F., CHINAPPEN-HORSLEY, U., MALKINA, I., RICHARDS, J. B., HAMMOND, N., STOLK, L., NICA, A., INOUYE, M., HOFMAN, A., STEPHENS, J., WHEELER, E., ARP, P., GWILLIAM, R., JHAMAI, P. M., POTTER, S., CHANEY, A., GHORI, M. J., RAVINDRARAJAH, R., ERMAKOV, S., ESTRADA, K., POLS, H. A., WILLIAMS, F. M., MCARDLE, W. L., VAN MEURS, J. B., LOOS, R. J., DERMITZAKIS, E. T., AHMADI, K. R., HART, D. J., OUWEHAND, W. H., WAREHAM, N. J., BARROSO, I., SANDHU, M. S., STRACHAN, D. P., LIVSHITS, G., SPECTOR, T. D., UITTERLINDEN, A. G. & DELOUKAS, P. 2009. Meta-analysis of genome-wide scans for human adult stature identifies novel Loci and associations with measures of skeletal frame size. *PLoS Genet,* 5**,** e1000445.

STRACHAN, D. P., WONG, H. J. & SPECTOR, T. D. 2001. Concordance and interrelationship of atopic diseases and markers of allergic sensitization among adult female twins. *J Allergy Clin Immunol,* 108**,** 901-7.

STRANGER, B. E., STAHL, E. A. & RAJ, T. 2011. Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics,* 187**,** 367-83.

SUGIMOTO, M., KAWAKAMI, M., ROBERT, M., SOGA, T. & TOMITA, M. 2012. Bioinformatics Tools for Mass Spectroscopy-Based Metabolomic Data Processing and Analysis. *Curr Bioinform,* 7**,** 96-108.

SUHRE, K., SHIN, S. Y., PETERSEN, A. K., MOHNEY, R. P., MEREDITH, D., WAGELE, B., ALTMAIER, E., CARDIOGRAM, DELOUKAS, P., ERDMANN, J., GRUNDBERG, E., HAMMOND, C. J., DE ANGELIS, M. H., KASTENMULLER, G., KOTTGEN, A., KRONENBERG, F., MANGINO, M., MEISINGER, C., MEITINGER, T., MEWES, H. W., MILBURN, M. V., PREHN, C., RAFFLER, J., RIED, J. S., ROMISCH-MARGL, W., SAMANI, N. J., SMALL, K. S., WICHMANN, H. E., ZHAI, G., ILLIG, T., SPECTOR, T. D., ADAMSKI, J., SORANZO, N. & GIEGER, C. 2011. Human metabolic individuality in biomedical and pharmaceutical research. *Nature,* 477**,** 54-60.

TORRES, M. J., BLANCA, M., FERNANDEZ, J., ROMANO, A., WECK, A., ABERER, W., BROCKOW, K., PICHLER, W. J. & DEMOLY, P. 2003. Diagnosis of immediate allergic reactions to beta-lactam antibiotics. *Allergy,* 58**,** 961-72.

TORRES, M. J., ROMANO, A., MAYORGA, C., MOYA, M. C., GUZMAN, A. E., RECHE, M., JUAREZ, C. & BLANCA, M. 2001. Diagnostic evaluation of a large group of

patients with immediate allergy to penicillins: the role of skin testing. *Allergy,* 56**,** 850-6.
TREUDLER, R. & SIMON, J. C. 2007. PPL and MDM skin test: new test kit is helpful in detecting immediate-type allergy to beta-lactams. *J Dtsch Dermatol Ges,* 5**,** 286-92.
UETRECHT, J. & NAISBITT, D. J. 2013. Idiosyncratic adverse drug reactions: current concepts. *Pharmacol Rev,* 65**,** 779-808.
VAJTR, D., SPRINGER, D., STANEK, L., BENADA, O., SAMAL, F., BERAN, M., STREJC, P., PRUSA, R. & ZIMA, T. 2014. Pathomorphology of inflammatory response following traumatic brain injury, serum values of interleukins, and gene polymorphisms. *Soud Lek,* 59**,** 40-7.
VISSCHER, P. M., BROWN, M. A., MCCARTHY, M. I. & YANG, J. 2012. Five years of GWAS discovery. *Am J Hum Genet,* 90**,** 7-24.
VOGLER, M. 2012. BCL2A1: the underdog in the BCL2 family. *Cell Death Differ,* 19**,** 67-74.
WEALE, M. E. 2010. Quality control for genome-wide association studies. *Methods Mol Biol,* 628**,** 341-72.
WELLCOME TRUST CASE CONTROL, C. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature,* 447**,** 661-78.
WILK, M. B. & GNANADESIKAN, R. 1968. Probability plotting methods for the analysis of data. *Biometrika,* 55**,** 1-17.
WOLD, A. E. 1998. The hygiene hypothesis revised: is the rising frequency of allergy due to changes in the intestinal flora? *Allergy,* 53**,** 20-5.
YANG, J., QIAO, H. L. & DONG, Z. M. 2005. Polymorphisms of IL-13 and IL-4-IL-13-SNPs in patients with penicillin allergies. *Eur J Clin Pharmacol,* 61**,** 803-9.
YANG, J., QIAO, H. L., ZHANG, Y. W., JIA, L. J., TIAN, X. & GAO, N. 2006. HLA-DRB genotype and specific IgE responses in patients with allergies to penicillins. *Chin Med J (Engl),* 119**,** 458-66.
ZHOU, X. & STEPHENS, M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet,* 44**,** 821-4.
ZUBERBIER, T., BINDSLEV-JENSEN, C., CANONICA, W., GRATTAN, C. E., GREAVES, M. W., HENZ, B. M., KAPP, A., KOZEL, M. M., MAURER, M., MERK, H. F., SCHAFER, T., SIMON, D., VENA, G. A., WEDI, B. & EAACI/GA2LEN/EDF 2006. EAACI/GA2LEN/EDF guideline: definition, classification and diagnosis of urticaria. *Allergy,* 61**,** 316-20.

Appendix A

Department of Twin Research & Genetic Epidemiology
St Thomas' Hospital Campus
4th Floor South Wing Block D
Westminster Bridge Road
London
SE1 7EH

KING'S College LONDON

University of London

Twins UK
Volunteer Information Sheet

1. Study Title

Twins UK

2. Invitation Paragraph

You are being invited to take part in a research study within the Department of Twin Research as you are currently registered on our twin database. Please take time to read the following information carefully and discuss it with friends, family or your GP if you wish to do so. Please ask us if there is anything that is not clear or if you would like to have more information. You will find a contact number at the end of this Information Sheet.

3. What is the purpose of this study?

The purpose of the study is to explore the genetic and environmental causes of a wide range of diseases and conditions. By using your DNA profile and the clinical and questionnaire data we can compare similarities and differences between identical (MZ) and non-identical (DZ) twin pairs to estimate the importance of genetic and environmental influences in a wide range of common conditions. This information can also be used to identify new genes and environmental factors associated with disease. Family members can also be used in the study to find out how diseases are influenced by genetic and environmental factors across the generations. Family members may be invited to participate.

4. Why have I been chosen?

You have been chosen to take part in this study as you are registered on our database. You may have previously had a visit at the Department of Twin Research. For this study we want to include all the twins (and those parents and sisters already registered) on our database and get up to date information on everybody.

5. Do I have to take part?

No, it is up to you to decide whether or not to take part in this study. If you decide that you do wish to participate, you will be given this information sheet to keep and we will ask you to sign a consent form. You will be given a copy of the consent form. You are free to withdraw at any stage

and this will not affect any future medical care or further research you may wish to participate in within the Department of Twin Research.

6. What will happen to me if I take part?

You will be invited to attend a visit at either the Department of Twin Research at St. Thomas' Hospital in London or one of our satellite centres throughout the UK, whichever is the most convenient for you to attend. During this visit we will collect saliva, buccal (mouth) cells, blood, urine, hair and nail samples from you and we will ask you to complete a series of self completed questionnaires.

You will undergo a variety of psychological and physiological tests which may include height, weight, blood pressure measurements, grip strength, photocopy of your hands, fingerprinting, facial photographs, electro-cardiogram (ECG), carotid ultrasound and measurement of the pulse using a probe, lung function test, distorted tune test for tone deafness, hearing tests, taste test, eye tests, a series of cognitive tests, bone mineral density scans and an exercise test involving walking and running on a treadmill at different gradients whilst having your heart rate measured. We may perform heat and pressure sensitivity tests using a heat probe on the forearm, brushing the skin with a brush or blunt probes, placing an extremity in an iced water bath. We may perform two 4mm skin biopsies on the abdomen under local anaesthetic and collect hair follicle samples for the purpose of stem cell research. This will enable us to investigate the way our genetic material determines cell function and how genetic malfunctions occur and cause disease. With the understanding of these genetic mechanisms we hope to find ways of overcoming or reversing these defective mechanisms. You may also be asked to take part in a trial involving taking a dietary supplement. We will inform you about the exact tests undertaken on your visit at the time of initial contact and you are free to opt out of any tests that you do not wish to have performed. With your express permission we may request additional information from your medical records.

7. What do I have to do?

All we ask is that you attend the unit having fasted overnight. This means no food or drink, except water, from midnight the night before your visit. You need to provide us with a generous urine sample, which has been **produced** 2 hours after you first went to the toilet. If you need to take any regular medication, please take them as you would normally do,

preferably with water only. **If you are diabetic, please do not fast but inform us of this prior to your visit.**
You will be required to complete a questionnaire prior to attending any of the centres. This will take about 30 minutes to fill in and may require you to find additional information, e.g. what drugs are you using now, in the past, and for how long have you used them.

8. What is the drug or procedure that is being tested?

Depending on the disease we are researching when you come for your visit, you may receive drugs such as eye drops or a local anaesthetic (numbing medicine) similar to that used at the dentist. We do not use any experimental drugs in any part of our study.

9. What are the side effects of taking part?

You should not experience any side effects whilst taking part in this study. You may feel a slight discomfort when the blood is being taken, but this should subside soon afterwards. The pain test might leave a small area of reddening but this will fade quickly after the test. The skin biopsy may be tender over a few days and may leave a slight scar.

None of the other tests performed such as height, weight, blood pressure measurements, grip strength, lung capacity test, eye tests, and carotid ultrasound will cause you any harm.

10. What are the possible disadvantages and risks of taking part?

There are no disadvantages and risks associated with this study. If you are pregnant, are planning to become pregnant or are of childbearing age, we may not perform certain tests on you, such as a Bone Mineral Density scan, or we may ask you to make the appointment for a later date. The bone mineral density scan delivers a small dose of radiation approximate to that received during a Trans-Atlantic Flight.

11. What are the possible benefits of taking part?

You will be informed of clinically relevant results of the tests performed on you during the visit. A copy of these results will also be sent to your GP.

12. What happens when the research study finishes?

Once the study has finished, your name will still remain on our database and you may be invited to participate in future projects. You will be

informed of further developments within the Department of Twin Research via regular annual Newsletters. You may also receive further questionnaires which will ask about any new diseases we are researching.

13. What if something goes wrong?

If you feel that you have been unfairly treated during your visit to any of the Centres, please direct your complaint to the Department of Twin Research who will investigate the nature of the complaint. You will receive feedback about your complaint within 21 working days. If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed due to someone's negligence, then you may have grounds for a legal action but you may have to pay for it.

14. Will my taking part in this study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. You will have been allocated a specific twin number and this will be used for all your data stored on our database. All information which leaves this department, except the results which are sent to you and your GP, will be anonymous, your name and address will have been removed and replaced by this unique study number, so that you can **not** be identified. All analyses are performed on large datasets, so no individual can be identified.

15. What will happen to the results of the research study?

The samples of blood, hair urine, nail, saliva, buccal (mouth) smear or tissue for DNA/RNA analysis and the information collected during your visit will be stored centrally or at collaborating establishments. They will be used for further research into common diseases and traits and will be accessible to independent researchers and collaborators who wish to use the enormous potential of the twin design to identify the action of individual genes, environmental exposures and their interaction by performing further tests on these samples. Your DNA will **not** be used for any experiments with regards to human cloning. Other academic centres and commercial companies, who are collaborating with us in our research, may request access and have rights to the data, including DNA and other samples, we have collected and to our research conclusions. The Department of Twin Research is now part of the Guy's and St Thomas' Hospital National Institute for Health Research (NIHR) Biomedical Resource Centre (BRC), a government funded initiative through the NHS to promote the sharing of knowledge, anonymous data and samples with

other leading scientific institutions. to support leading edge research focused on the needs of the patient and public.

The results of the genetic research, such as the types of genes you have or the sequence of your DNA, may be stored in an electronic archive (the European Genome-Phenome Archive at the European Bioinformatics Institute, Hinxton, Cambridge). This data will always be stored in an anonymised form but may be made accessible to *bona fide* researchers at other research establishments *via* a managed access system.

The data/intellectual property collected as a result of your participation will always reside with the Hospital and Kings College London. We often rely on the expertise of commercial companies to share data for writing of papers and to advance our research. In the rare event of any commercial profits, funds will come back to the hospital and the university for the further advancement of research.

You will not have claim to either the material that we collected or the result arising from the study. If you decide to withdraw your future consent for biological sample collection, previous samples cannot be removed or destroyed and will continue to be used for future research. We shall ensure that all personal data and results are anonymised and cannot lead to the identification of a single individual.

16. Who is organising and funding the research?

The study is organised by the Department of Twin Research at St. Thomas' Hospital in London and at present funding for this research project has been provided by the Wellcome Trust, a registered Charity and the Biomedical Resource Centre.

17. Who has reviewed the study?

The study has been reviewed and approved by the St. Thomas' Hospital Ethics Committee.

18. Contact for further information

If you require any further information about this study or would like some questions answered, please do not hesitate to contact the Department of Twin Research on Tel. 020 7188 5555 Monday to Friday between 9am to 5pm.

Thank you for taking the time to read this Information Sheet

Appendix B

Department Twin Research & Genetic Epidemiology
St Thomas' Hospital Campus
4th Floor South Wing Block D
Westminster Bridge Road
London SE1 7EH

KING'S College LONDON

University of London

Centre Number:
Study Number: **EC04/015**
Patient Identification Number for this trial:

CONSENT FORM

Title of Project: TwinsUK

Name of Researcher: **Prof Tim Spector**

		Please Initial ↓
1	I confirm that I have read and understood the information sheet version **7** dated **04.04.12** for the above study and have had the opportunity to ask questions.	
2	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	
3	I understand that relevant sections of my medical notes, health related records and data collected during TwinsUK, may be looked at by responsible individuals from collaborating companies, regulatory authorities, the NHS Trust and the NIHR Bioresource, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records.	
4	I agree that the samples I have donated and the information gathered about me can be sharedwith the NIHR Bioresource and *bone fide* researchers and stored for use in future research studies aimed at identifying the interactions between genes, the environment and disease.	
5	I agree to my GP being informed of my participation in the study and to receiving my clinically relevant results.	
6	I agree to have the assessments stated in the information sheet performed and consent to take part in the above study.	

Name of Participant (BLOCK CAPITALS)	Date	Signature
Name of Person taking consent (if different from researcher)	Date	Signature
Researcher	Date	Signature

When completed, 1 for patient, 1 for researcher site file, 1 (original) to be kept in medical notes

Appendix C

S1 – Peas in the Pod Questionnaire (PPQ)

ZYGOSITY

Dear Twin,

For research purposes we ask all twins to complete this short questionnaire so we can collect further information from which to ascertain your zygosity (i.e. whether you are an identical or non-identical twin), as some twins are uncertain about this. It is very important that you DO NOT confer with your twin otherwise it will give a false result. Please could you kindly complete the questions below and return with your questionnaire, even if you are sure about your zygosity.

Many thanks**.**

_____

1. a) At school, did people have trouble telling you apart?

(0) O Yes

(2) O No

(1) O I don't know

b) Were your parents able to tell you apart at school age?

(2) O Yes

(0) O No

(1) O I don't know

c) Were your close school friends able to tell you apart at school age?

(2) O Yes

(0) O No

(1) O I don't know

d) Were strangers able to tell you apart at school age?

(2) O Yes

(0) O No

(1) O I don't know

e) In childhood, which of the following would best describe you and your twin? (Please select one)

(0) O As alike as peas in a pod
(2) O Ordinary sibling likeness (like sisters or brothers)
(1) O I don't know

NOTE: We will contact you for further information if your Zygosity on completing this Questionnaire is different from your Self Report

Appendix D

Dr TD Spector
Twin Research Unit
Block 4A, South Wing
St Thomas' Hospital

ST THOMAS' HOSPITAL
RESEARCH ETHICS COMMITTEE
Ethics Committee Office
Block 5, South Wing,
St Thomas' Hospital
London SE1 7EH

Chairman – Dr A Hopper/Dr AJ Williams
Administrator - Ms S Hirsch

Phone: 020 7928 9292 Ext 2097
Fax: 020 7922 8163
Email: Stella.hirsch@gstt.sthames.nhs.uk

15 March 2004

Dear Dr Spector

EC04/015 Twins UK *Dr TD Spector, Prof Alex MacGregor*

- Local Consent Form Version no 1, 01/02/04,
- Consent form for DNA studies Version 1 date 10/03/04
- Volunteer information sheet Version 1, 01/02/2004

Thank you for addressing the queries raised by the Research Ethics Committee at its meeting on 24 February 2004. This is satisfactory and I am happy for the study to commence. Approval extends to the Guy's site.

Please note the following conditions to the approval:

- You do not undertake this research in a NHS organisation until the relevant NHS management approval has been gained (R&D).
- The project number and the principal investigator must be clearly stated on the consent form (if applicable). If approval is given to named investigators only, these names must also be stated on the form.
- In the case of research on patients, a copy of the consent form (if applicable) must be placed in the patient's medical records, together with a note of the date of commencement of his/her participation in the research. A label must appear on the outside cover of the records when the patient is participating in the research.
- Any amendments to the protocol must be notified to the committee for approval.
- Approval is for the length of time specified in your application. If you require an extension, a letter from the principal investigator to the Chairman, is required to extend the research.
- The committee should be notified of any serious adverse events (please apply for standard SAE report form), or if the study is terminated prematurely
- The investigators must adhere to the published Guidelines of the Committee and provide the Chairman with annual progress reports and an end of study report. The research should start within 12 months of the date of approval.

This project carries a reference number, noted above, which must be quoted in any future correspondence.

The St Thomas' Hospital LREC is compliant with the ICH GCP requirements.

Yours sincerely

Dr AJ Williams
Co-Chairman
Research Ethics Committee

Encl.

Appendix E

NHS
Health Research Authority

NRES Committee London - Westminster

(Formerly St Thomas' Ethics Committee)
Research Ethics Committee (REC) Centre, Charing Cross,
Room 12, 4th Floor West,Charing Cross Hospital
Fulham Palace Road,
London
W6 8RF

Tel: 020 3311 0100
Fax: 020 3311 7280

14 May 2012

Professor Tim Spector
Director of Dept. of Twin Research & Genetic Epidemiology
Kings College London University
STH, 1st Flr, South Wing, Block 4a
Westminster Bridge Road
London
SE1 7EH

Dear Professor Spector

Study title:	**Twins UK**
REC reference:	**EC04/015**
Amendment number:	**2**
Amendment date:	**04 April 2012**

The above amendment was reviewed at the meeting of the Sub-Committee held on 20 April 2012.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Participant Consent Form	8	04 April 2012
Participant Information Sheet: Volunteer Information Sheet	7	04 April 2012
Notice of Substantial Amendment (non-CTIMPs)	2	04 April 2012
Covering Letter		04 April 2012

Membership of the Committee

The members of the Committee who took part in the review are listed on the attached sheet.

R&D approval

All investigators and research collaborators in the NHS should notify the R&D office for the relevant NHS care organisation of this amendment and check whether it affects R&D approval of the research.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

EC04/015:	**Please quote this number on all correspondence**

Yours sincerely

P.P.

Dr L Alan Ruben
Chair

E-mail: Laura.Keegan@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

NRES Committee London - Westminster

Attendance at Sub-Committee of the REC meeting on 20 April 2012

Name	*Profession*	*Capacity*
Dr Anthony Kaiser	Consultant Neonatologist	Expert
Dr L Alan Ruben	Retired General Practitioner	Expert
Miss Ros Stanbury	Ophthalmologist	Expert

Appendix F

DRUG HYPERSENSITIVITY

Protocol No: .....
Date of protocol: .....

INVESTIGATOR:
Name:.....Center:.....

Address:.....Tel/Fax/E-mail: .....

PATIENT:
Name:.....Date of birth:..... Age:.....years
Weight:.....kg Height:.....cm
Profession:.....Origin:.....Sex: ☐ M ☐ F
Risk groups: ☐ Medical staff ☐ Pharmaceutical industries ☐ Farmers ☐ Others/specify.....

CURRENT COMPLAINTS:.....
.....
.....
.....

DRUG REACTION: **DATE OF REACTION:**.....
(Multiple boxes can be ticked; underline the choice if necessary; chronology can be characterized with numbers)

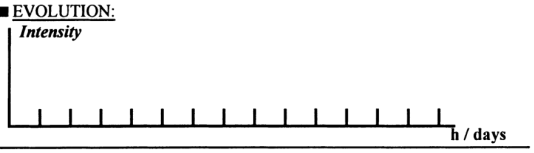
■ CUTANEOUS SYMPTOMS:
☐ Maculopapular exanthema
☐ Macular exanthema
☐ Urticarious exanthema
☐ AGEP (Acute generalized exanthemous pustulosis)
☐ Eczematoid exanthema
☐ Erythema exudativum multiforme
☐ Bullous exanthema
☐ Stevens-Johnson syndrome / TEN (M. Lyell)
☐ Fixed drug exanthema
☐ Purpura -> Thrombocyte count :.....
☐ palpable ☐ haemorrhagic-necrotizing
☐ Visceral organ involvement:.....
☐ Contact dermatitis ☐ Topic cause ☐ Haematogenous cause ☐ .....
☐ Urticaria Vasculitis
☐ **ONLY** pruritus
☐ Urticaria
☐ Angioedema/Location/s: .....
☐ Conjunctivitis
☐ Other/Specification:.....

☐ Morphology/Location/s:.....

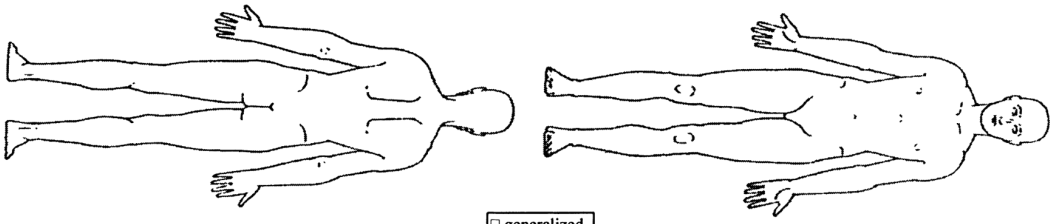
.....

■ DIFFERENTIAL DIAGNOSIS:
☐.....
☐.....
☐.....

■ CONTRIBUTING FACTORS:
☐ Viral infections: ☐ Flu-like infection ☐ Other:.....
☐ Fever
☐ Suspicion of photosensitivity ? ☐ No ☐ Yes ☐ Unknown
☐ Stress
☐ Exercise
☐ Other/Specification: .....
.....

■ EVOLUTION:


■ EFFLORESCENCES: Distribution / Dynamics (⇑ ⇓)



☐ generalized

■ GASTROINTESTINAL AND RESPIRATORY SYMPTOMS:
☐ Nausea/Emesis
☐ Diarrhea
☐ Gastrointestinal cramps

☐ Cough
☐ Dysphonia
☐ Dyspnea PEFR or FEV_1 :.....
☐ Wheezing/Bronchospasm

☐ Rhinitis
☐ Rhinorrhea
☐ Sneezing
☐ Nasal obstruction
☐ Other/Specification:.....

■ PSYCHIC SYMPTOMS:
☐ Fear/Panic reaction ☐ Vertigo
☐ Fainting
☐ Paraesthesia/Hyperventilation
☐ Sweating
☐ Other/Specification:.....

■ ASSOCIATED SYMPTOMS:
☐ Involvement of: ☐ Liver ☐ Kidney ☐ Other/Specification: .....
☐ Fever .....°C
☐ Malaise
☐ Pain/Burning ☐ Location/s:.....
☐ Edema ☐ Location/s:.....
☐ Arthralgia/Myalgia ☐ Location/s:.....
☐ Lymphadenopathy
☐ Other/Specification:.....

■ CARDIOVASCULAR SYMPTOMS:
☐ Tachycardia Pulse rate: ...../min
☐ Hypotension Blood pressure: .....mmHg
☐ Collapse
☐ Arrhythmia
☐ Other/Specification: .....

■ INVOLVEMENT OF OTHER ORGANS :
(e.g. peripheral neuropathy, lung involvement, cytopenia, etc.)
☐.....
☐.....
☐.....

■ **CLINICAL OUTCOME:** .....
.....
.....

■ List all drugs including Over The Counter substances, natural remedies and additive-containing food taken at the time of the reaction:
.....
.....
.....
.....

■ **SUSPICIOUS DRUGS:**

Drug's generic name ± additives / Indication:	Daily dose / Route of application / Duration of therapy:	Interval between dose and reaction	Previous therapy with this drug:
1.	.....mg/d; .....; .....d		☐ No ☐ Unknown ☐ Yes ->
Symptoms:.....
2.	.....mg/d; .....; .....d		☐ No ☐ Unknown ☐ Yes ->
Symptoms:.....
3.	.....mg/d; .....; .....d		☐ No ☐ Unknown ☐ Yes ->
Symptoms:.....
4.	.....mg/d; .....; .....d		☐ No ☐ Unknown ☐ Yes ->
Symptoms:.....
5.	.....mg/d; .....; .....d		☐ No ☐ Unknown ☐ Yes ->
Symptoms:.....
6.	.....mg/d; .....; .....d		☐ No ☐ Unknown ☐ Yes ->
Symptoms:.....

.....

☐ CURRENT DRUGS: ..... ☐ Antihistamines .....
☐ β-Blockers .....

■ MANAGEMENT FOLLOWING ACUTE DRUG REACTION: ☐ No therapy
☐ Stopping suspicious drugs No.# .....
☐ Antihistamines ☐ local ☐ systemic
☐ Corticosteroids ☐ local ☐ systemic
☐ Bronchodilatators ☐ local ☐ systemic
☐ Shock treatment ☐ Epinephrine ☐ Plasma expanders ☐ Other: .....
☐ Change to substitute/s:
☐ Type/Name: .....
☐ Tolerance: .....
☐ Other/Specification:.....
☐ Dosis reduction (Drug.....).....
☐ Other/specify.....
.....

PERSONAL HISTORY:

1) HAVE SIMILIAR SYMPTOMS BEEN OBSERVED WITHOUT THE INTAKE OF THE SUSPICIOUS DRUGS ? ☐ Yes ☐ No ☐ Unknown

2) MEDICAL HISTORY:
☐ Asthma ☐ Autoimmune (Sjögren, Lupus, etc.) ☐ Urticaria pigmentosa / syst. mastocytosis
☐ Nasal polyposis ☐ Lymphoprolific (ALL, CLL, Hodgkin, etc.) ☐ Chronic urticaria
☐ Cystic fibrosis ☐ Intervertebral disk surgery ☐ HIV positivity
☐ Diabetes ☐ Liver:..... ☐ Kidney: .....
☐ Other/Specification: .....

3) ALLERGIC DISEASES: .....
(e.g. pollinosis, atopic dermatitis, food allergy, Hymenoptera venom allergy, latex allergy, etc.)

4) DRUG REACTIONS DURING PREVIOUS SURGERY:.....☐ Dentist ☐ Local anaesthesia ☐ General anaesthesia (No:.....)

5) REACTIONS DURING PREVIOUS VACCINATIONS:.....☐ Polio ☐ Tetanus ☐ Rubella ☐ Measles ☐ Hepatitis B
☐ Diphtheria ☐ Other:..... ☐ Unknown

FAMILY HISTORY: Allergies / Drug allergies:
.....

REMARKS:
.....
.....
.....

DIAGNOSTIC PROCEDURES:

RESULTS

■ ACUTE DIAGNOSTICS: (already performed) ..... DATE

		NORMAL	ABNORMAL	QUESTIONABLE
☐ Blood:	☐ Full blood count: ☐ Eosinophils:	☐	☐ Value:.....rel.; .....abs.	☐
	☐ Other:.....	☐	☐ Value:.....rel.; .....abs.	☐
	☐ ECP (Eosinophil cationic protein)	☐	☐ Value:.....	☐
	☐ C-reactive protein / Erythrocyte sedimentation rate	☐	☐ Value:.....	☐
	☐ Flowcytometry (Specify: .....)	☐	☐.....	☐
	☐ Tryptase	☐	☐ Value:.....	☐
☐ Liver parameters:	☐ GOT	☐	☐ Value:.....	☐
	☐ GPT	☐	☐ Value:.....	☐
	☐ γGT	☐	☐ Value:.....	☐
	☐ alk. Phosphatase	☐	☐ Value:.....	☐
☐ Kidney:	☐ Creatinine	☐	☐ Value:.....	☐
	☐ Methylhistamine	☐	☐ Value:.....	☐
	☐ Other:.....	☐	☐.....	☐
☐ Special:	☐ Mediators and metabolites (IL-4, IL-5, IL-10, IFNγ)	☐	☐ Value:.....	☐
	☐ Immune complex analysis	☐	☐	☐
	☐ Complement analysis	☐	☐	☐

☐ Skin biopsy: .....

■ DIAGNOSTICS:

		NEGATIVE	POSITIVE	QUESTIONABLE
Skin tests:	☐ Prick :	☐	☐ Immediate-R . ☐ Late-R.	☐
	.....	☐	☐ Immediate-R. ☐ Late-R.	☐
	.....	☐	☐ Immediate-R. ☐ Late-R.	☐
	☐ Intradermal:	☐	☐ Immediate-R. ☐ Late-R.	☐
	.....	☐	☐ Immediate-R. ☐ Late-R.	☐
	☐ Scratch-Patch:	☐	☐ Immediate-R. ☐ Late-R.	☐
	.....	☐	☐ Immediate-R. ☐ Late-R.	☐
	.....	☐	☐ Immediate-R. ☐ Late-R.	☐
	☐ Other:	☐	☐ Immediate-R. ☐ Late-R.	☐

Blood analysis:	☐ Total IgE	☐ Value:.....	
	☐ Specific IgE for drugs: ☐ CAP ☐ RAST	☐ Value:.....	
	.....	☐ Value:.....	
	.....	☐ Value:.....	
	.....	☐ Value:.....	
	☐ Specific IgG/Coombs Test dir:		
	☐ Coombs test indir.		
	☐ Other.....		

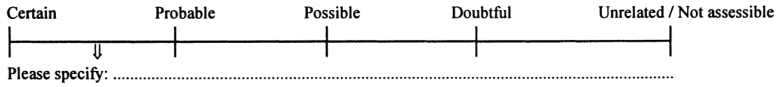
Cellular tests:	☐ Lymphocyte transformation test (LTT):.....	☐	☐ SI:.....	☐
	.....	☐	☐ SI:.....	☐
	.....	☐	☐ SI:.....	☐
	☐ Basophil activation test (Specify: .....)	☐	☐	
	☐ CAST assay	☐	☐	☐
	☐ Other:.....	☐	☐	

Provocation tests:	☐ Local anaesthetics:.....	☐	☐
	☐ NSAID:.....	☐	☐
	.....	☐	☐
	☐ Aspirin	☐	☐
	☐ Paracetamol	☐	☐
	☐ Nimesulid	☐	☐
	☐ β-lactam antibiotics:.....	☐	☐
	.....	☐	☐
	☐ Other:.....	☐	☐

CONCLUDING INTERPRETATION:

☐ Type I reaction (IgE mediated)	to: A.....
☐ Type II reaction (antibody mediated)	to: B.....
☐ Type III reaction (immune complexe mediated)	to: C.....
☐ Type IV reaction (cell-mediated, late-type reaction)	to: D.....
☐ Cytotoxic reaction, cell-mediated	to: E.....
☐ Pseudoallergic reaction	to: F.....
☐ Pharmacological reaction	to: G.....
☐ Psychophysiological reaction	to: H.....
☐ Other:.....	to: I.....

■ PROBABILITY SCALE CONCERNING THE CAUSAL RELATIONSHIP BETWEEN DRUG & REACTION:
(Please mark the drug's letter on the scale)



Please specify: .....

☐ DECLARATION TO REGULATORY AGENCY ?: ☐ No ☐ Yes ☐ To whom ?: .....Date: .....

REMARKS:.....

.....

Appendix G

Dr. Kourosh R Ahmadi PhD
St Thomas' Hospital
Dept Twin Research & Genetic Epidemiology
Lambeth Palace Rd
London SE1 7EH
Tel: 020 7188 6728
Email: kourosh.ahmadi@kcl.ac.uk

Dr. Rosario Caballero MD PhD
Guy's Hospital
Dept of Allergy & Respiratory Science
St Thomas Street
London SE1 9RT
Tel: 020 7188 5075
Email: rosario.caballero@kcl.ac.uk

KING'S College LONDON
University of London

PARTICIPANT INFORMATION SHEET

LREC reference: 11-LO-0112

Study title: Genetic Epidemiology of Immediate Allergy to Beta-Lactam Antibiotics

You are being invited to take part in a research study. Before you decide to take part, it is important for you to understand why the research is being done and what it will involve. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Beta-lactam antibiotics, including penicillin and cephalosporins, are the most widely used antibiotics to treat common infections. Unfortunately, in up to 10% of the people they cause a marked, unwanted, and immediate allergic reaction. Such adverse reactions not only represent as a significant burden to the patient but also increase the overall cost of treatment markedly. Recent studies have shown that some people face greater risk of suffering such adverse events which can be due to a combination of factors including genetic, environmental exposure (diet, infection, interactions with other drugs), as well as other epidemiological (age, sex) aspects.

The purpose of this study is to uncover genetic and environmental factors that increase the risk of allergy to beta-lactam antibiotics. More specifically, we will attempt to identify genetic variations – mutations – as well as environmental or epidemiological factors that are associated with the immediate allergic response to beta-lactam antibiotics. The results will improve clinical prediction of immediate allergic responses to these commonly prescribed drugs and so reduce the incidence of these unwanted responses.

Why have I been invited?

You have been invited to take part in this study because you have previously had an immediate reaction following administration of beta-lactam antibiotics. If you have not had an immediate reaction to one of these antibiotics, you have been invited to join the study as a control participant.

Do I have to take part?

It is totally up to you to decide whether or not to take part. If you agree to participate in the study you will be given this information sheet to keep and asked to sign a consent form (attached). If you do decide to take part you are still free to withdraw at any time and without having to give a reason. Whether you decide not to take part or initially agree to take part and later change your mind, the standard of your care in hospital, either presently or at any time in the future, will not be affected in any way.

What will happen to me if I take part?

If this is your first visit to the Allergy clinic and you have not previously had a beta-lactam allergy test, and agree to participate in this study, you will be asked to:

- Sign a consent form.
- Undergo a clinical examination. During this examination we will check your vital signs including pulse rate, blood pressure and blood oxygen (using a finger clip).
- Have skin tests to see whether or not you are allergic.
  In these tests, we start with skin prick tests, followed by intra-dermal tests. Skin prick tests are done placing on the skin of your forearm a small drop of 5 different beta-lactam drugs (benzylpenicillin, benzylpenicilloyl poly-L-lisine, minor determinant mixture, amoxicillin, and other beta-lactam antibiotics when necessary). Fluid from these drops are introduced onto the surface of the skin by pricking with a sterile aluminium lancet. The skin prick is mildly painful and very rapid. If you are allergic to a particular drug, the site forms a small itchy blister (positive result). If you are not allergic, nothing happens and we carry on with intra-dermal tests. Intra-dermal tests are done by injecting 0.05ml (a tiny volume like a splash of water) of the same 5 beta-lactam drugs into our forearm. If you have a positive result to any beta-lactam antibiotics at any stage, then skin testing with that specific antibiotic is stopped. The skin testing procedure takes 1 hour. Any positive itchy reaction will settle and resolve within 10 mins (see section "What are the side effects of taking part").
- Give a sample of blood of 20 ml (about 2 tablespoonfuls) which will be used to extract DNA, for carrying out genetic profiling, or to carry out further biochemical tests useful for characterising the allergy.

If you have been seen in the Allergy clinic previously and have already been diagnosed as allergic to beta-lactam antibiotics; if you agree to take part in this study you will be:

- Asked to visit the clinical research facilities at Guy's Hospital.
- At the visit, we will initially answer any outstanding questions you may have about the study;
- If you are still happy to participate, we will ask you to sign a consent form which will allow you to participate in the study;
- We will collect 20 ml of blood (about 2 tablespoonfuls) which will be used to extract DNA, for carrying out genetic profiling, or to carry out further biochemical tests useful for characterising the allergy.

How long will the visits last?

The visit should take no longer that 1 hour.

What are the side effects of taking part?

Some people find skin testing slightly uncomfortable. As mentioned above, if you are allergic to a particular drug, the test site forms a small itchy blister. This can be itchy for about 10 minutes. Most people tolerate this very well but if necessary a dose of antihistamine will be given after the testing to minimise the discomfort.

Skin testing carries a small risk of provoking a more widespread allergic reaction, with generalised itchy skin rash, chest tightness and low blood pressure. This is extremely uncommon. In the rare case in which it occurs, it happens within 30 minutes of injection of the allergen, which is why we ask you to remain in the clinic for half an hour. It is readily and rapidly treatable and we are experienced in doing this. Skin testing will have been performed as part of your routine drug allergy clinic assessment.

Finally, you may feel a slight discomfort when blood is being taken, but this should subside soon afterwards.

What are the possible benefits of taking part?

The results of this study should help identify individuals who have a greater risk of beta-lactam antibiotic allergy and so inform medical practitioners in their choice of antibiotic drug prescription.

What if I want to withdraw from the study?

You may withdraw from the study at any time. You do not have to give a reason. Withdrawing from the study will not influence your current or future treatment in any way. Should you wish to withdraw from the study after giving us your blood sample, we will destroy any blood or DNA and so will not be able to use these samples further in any of our investigations

Will my participating in this study be kept confidential?

All information which is collected about you during the course of this research will be kept strictly confidential. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised from it.

Storage of your blood sample

As part of the studya blood sample will be collected as described earlier. We ask for your permission to store any surplus blood or DNA extracted from the blood in our laboratories for future use in other genetic studies into allergy. The samples will be identifiable as having come from you, but will be used only for the purpose of research and not for any diagnostic purpose. No personal information will be given to any third party without your written consent. Samples will be stored for no more than 10 years after which left over samples will be destroyed. Please note that you can withdraw your consent for us to store your samples at any time, when they would then be destroyed and this will communicated back to you by a letter.

What will happen to the results of this research study?

Eventually the results of the study will be published in a medical journal. You will be welcome to receive from us any resulting publications if you so wish, but the study results will be entirely anonymous and you will not be able to identify your individual results.

Who is organising and funding the research?

The study is being funded in part by Biomedical Research Centre maintenance funding. Guy's and St Thomas' NHS Foundation Trust, along with its academic partner King's College London, is one of five new comprehensive Biomedical Research Centres in the UK funded by the National Institute for Health Research. These centres have a strong focus on translational research, taking advances in basic medical research out of the laboratory and into the clinical setting so that they can benefit patients. They form a key part of the Department of Health's new strategy for research and development in the NHS.

Who has reviewed the study?

The study was reviewed and approved by the National Research Ethics Service, South East London REC 2.

Contact for further information

For further information about this study, please contact Dr Kourosh Ahmadi (email: kourosh.ahmadi@kcl.ac.uk or telephone on 020 7188 6728) or Dr Rosario Caballero (email: rosario.caballero@kcl.ac.uk or telephone on 020 7188 5075).

Thank you for your time in reading this information sheet and considering participating in this research study.

Please keep this information sheet and a copy of your consent form for future reference.

Appendix H

Table 1: Details of metabolites measured in the Metabolon platform

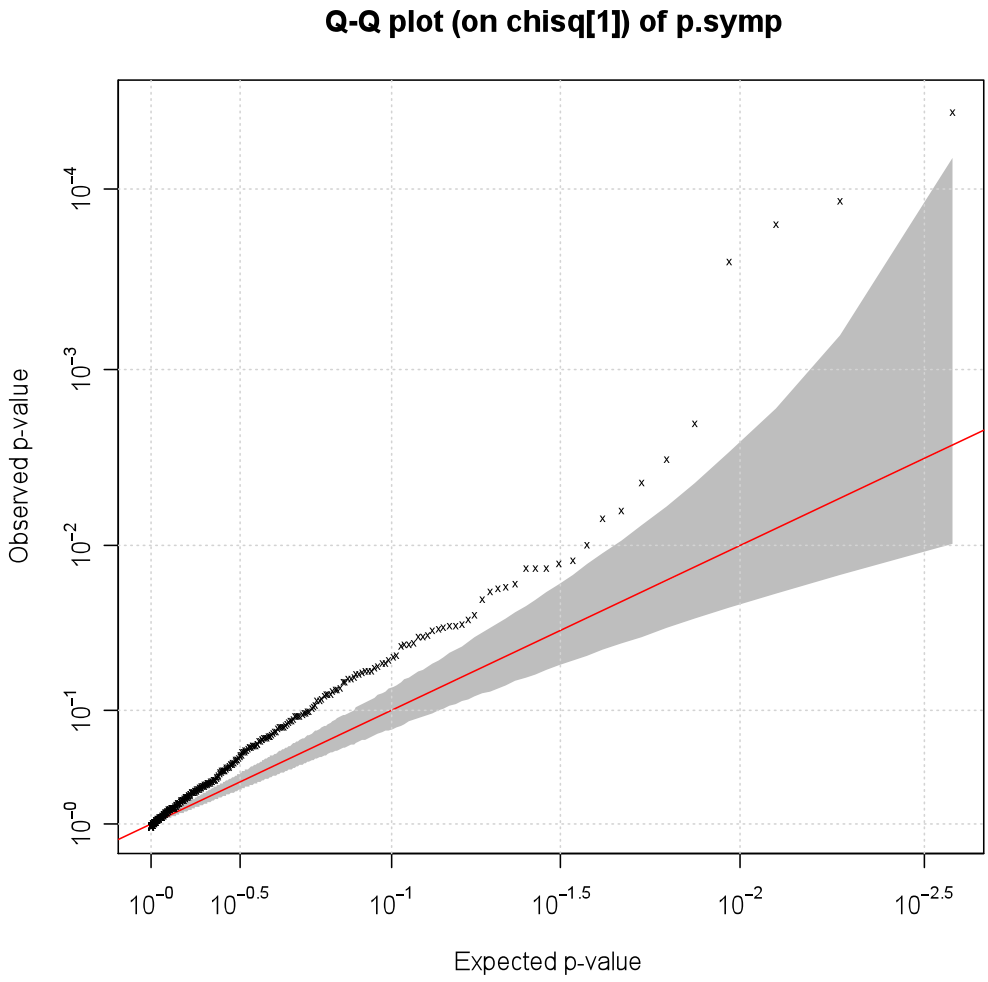
Metabolite group	*n*	Sub-pathways	Examples
Amino acids	70	Alanine and aspartate metabolism; Butanoate metabolism; Creatine metabolism; Cysteine, methionine, SAM, taurine metabolism; Glutamate metabolism; Glycine, serine and threonine metabolism; Guanidino and acetamido metabolism; Histidine metabolism; Lysine metabolism; Phenylalanine and tyrosine metabolism; Tryptophan metabolism; Urea cycle – arginine and proline metabolism; Valine, leucine and isoleucine metabolism	Alanine, arginine, beta-hydroxyisovalerate, indolelactate, N-acetylalanine, trans-4-hydroxyproline
Carbohydrates	13	Aminosugars metabolism; Fructose, mannose, galactose, starch, and sucrose metabolism; Glycolysis, gluconeogenesis, pyruvate metabolism; Nucleotide sugars, pentose metabolism	Glucose, fructose, arabinose, mannitol, pyruvate
Cofactors and vitamins	11	Ascorbate and aldarate metabolism; Hemoglobin and porphyrin metabolism; Pantothenate and CoA metabolism; Tocopherol metabolism; Vitamin B6 metabolism	alpha-tocopherol, biliverdin, threonate
Energy metabolism	6	Citric acid (Krebs) cycle; Oxidative phosphorylation	Citrate, acetylphosphate
Lipids	122	Fatty acids (essential, branched, dicarboxylate, ester, monohydroxy, short chain, medium chain, long chain); Ketone bodies; Lysolipids; Monoacylglycerols; Sphingolipids; Sterols/Steroids; Bile acid metabolism; Carnitine metabolism; Eicosanoids; Fatty acid metabolism; Fatty acid metabolism; Inositol metabolism	1-oleoylglycerol, 1-oleoylglycerophosphocholine, acetylcarnitine, androsterone sulphate, chiro-inositol, cortisol, deoxycholate, DHEAS, thromboxane B2
Nucleotides	11	Purine metabolism; Pyrimidine metabolism	Guanosine, pseudouridine, xanthine
Peptides	21	Dipeptides; Polypeptides; Fibrinogen cleavage peptides; gamma-glutamyl peptides	Aspartylphenylalanine, gamma-glutamylleucine
Xenobiotics	45	Drugs; Chemicals; Food components; Tobacco metabolites; Sugar, sugar substitute, starch; Benzoate metabolism; Xanthine metabolism	Atenolol, caffeine, cotinine, erythritol, ibuprofen, paraxanthine, saccharin
Unknown	211	-	-
Total	**510**		

Appendix I

Results

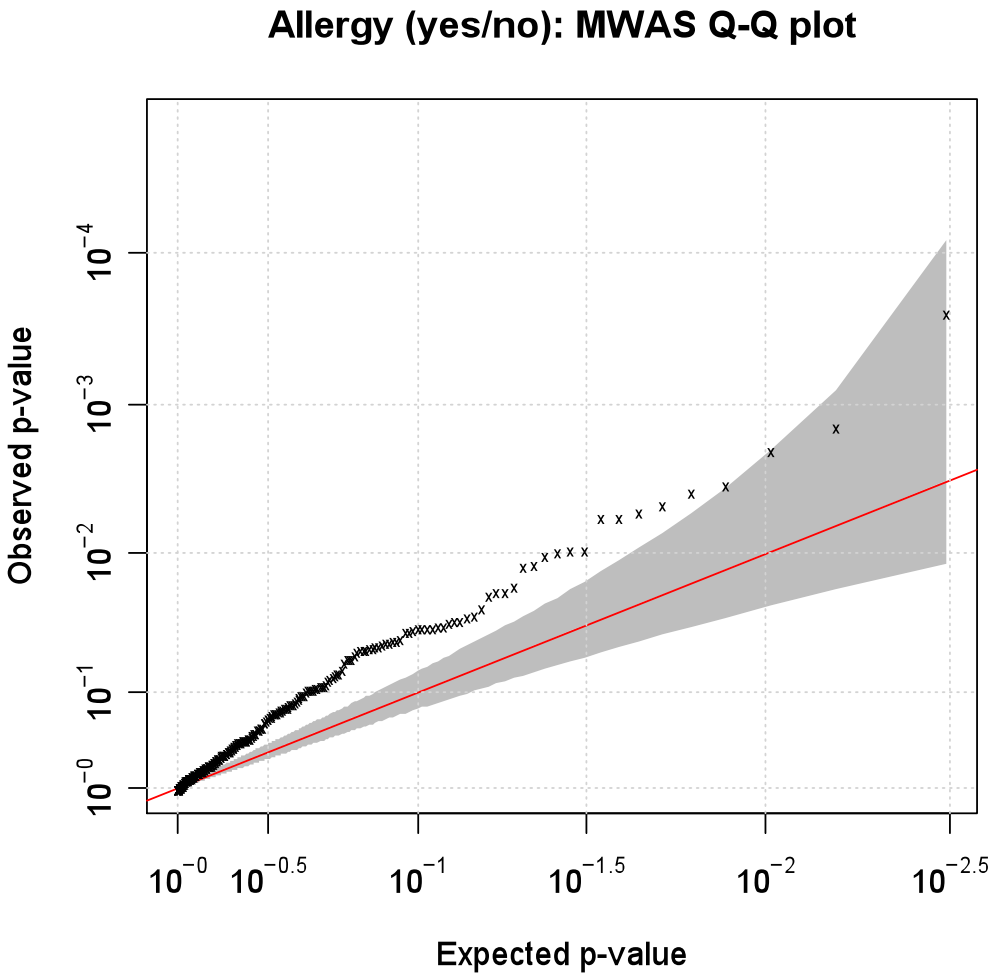
Test 1:

Q-Q plot



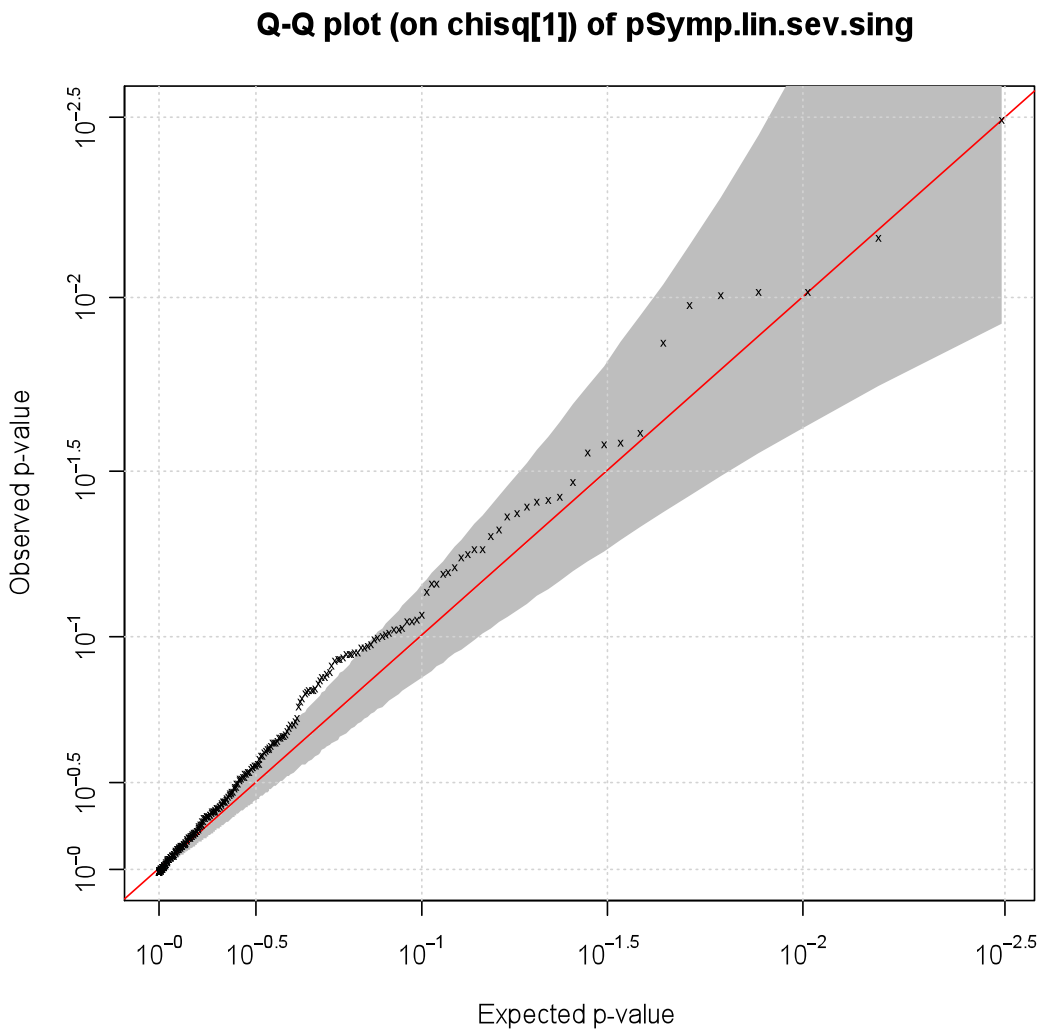
Test 2:

Q-Q plot



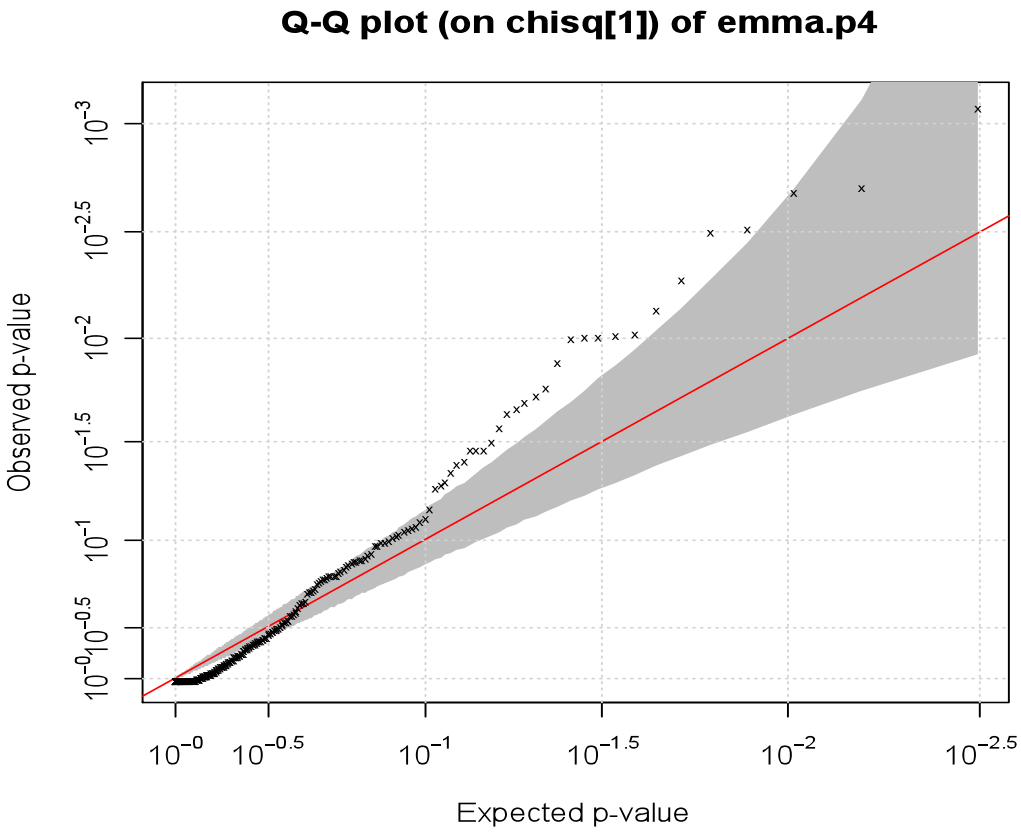
Test 3:

Q-Q plot



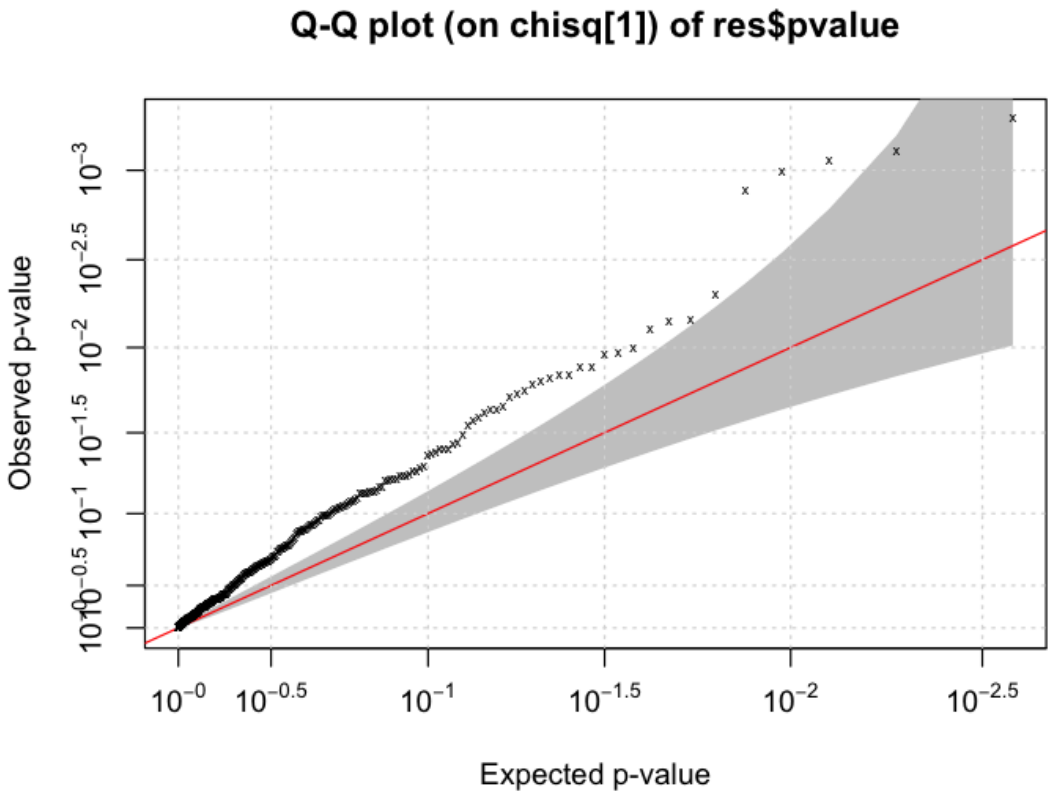
Test 4:

Q-Q plot



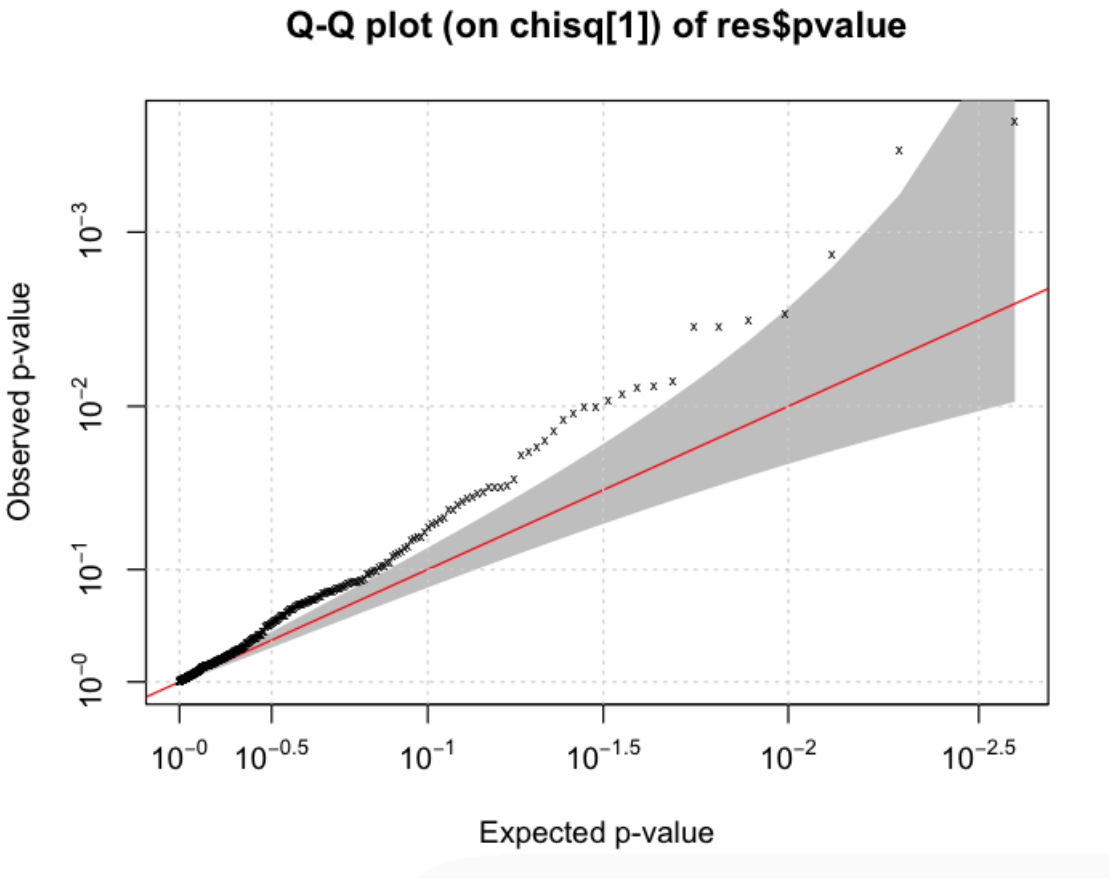
Test 5:

QQ-plot



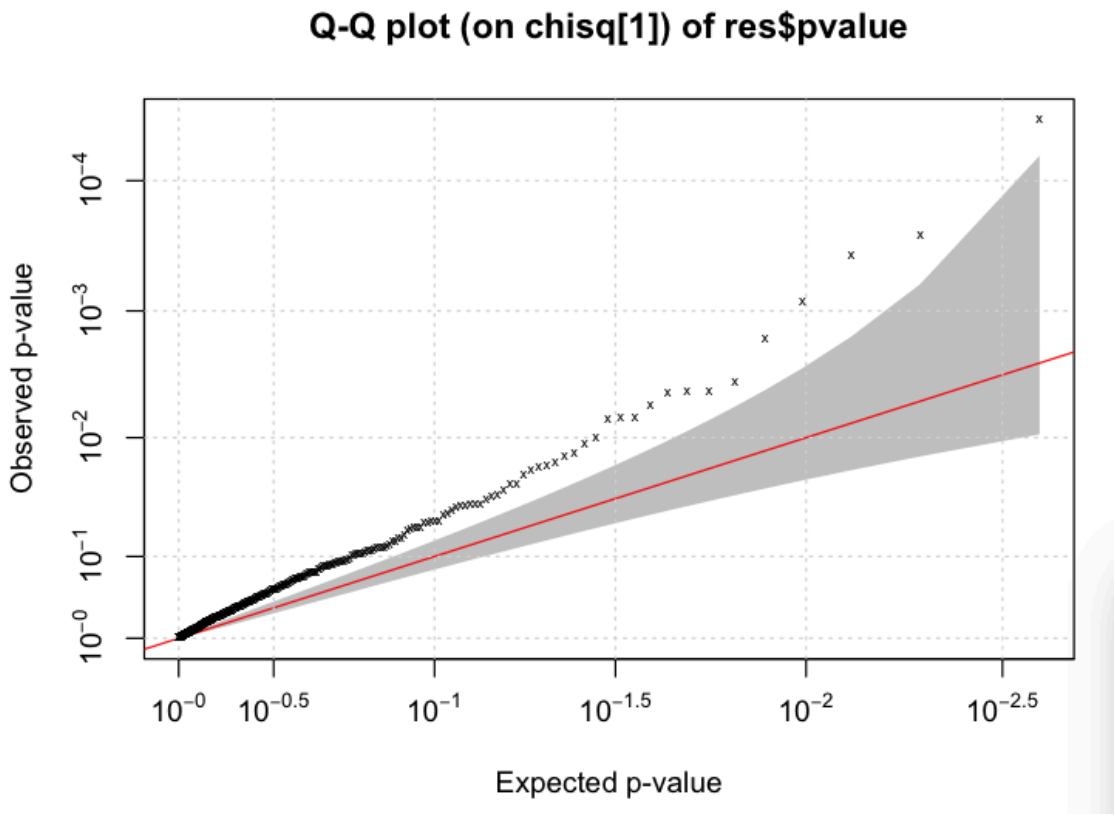
Test 6:

Q-Q plot



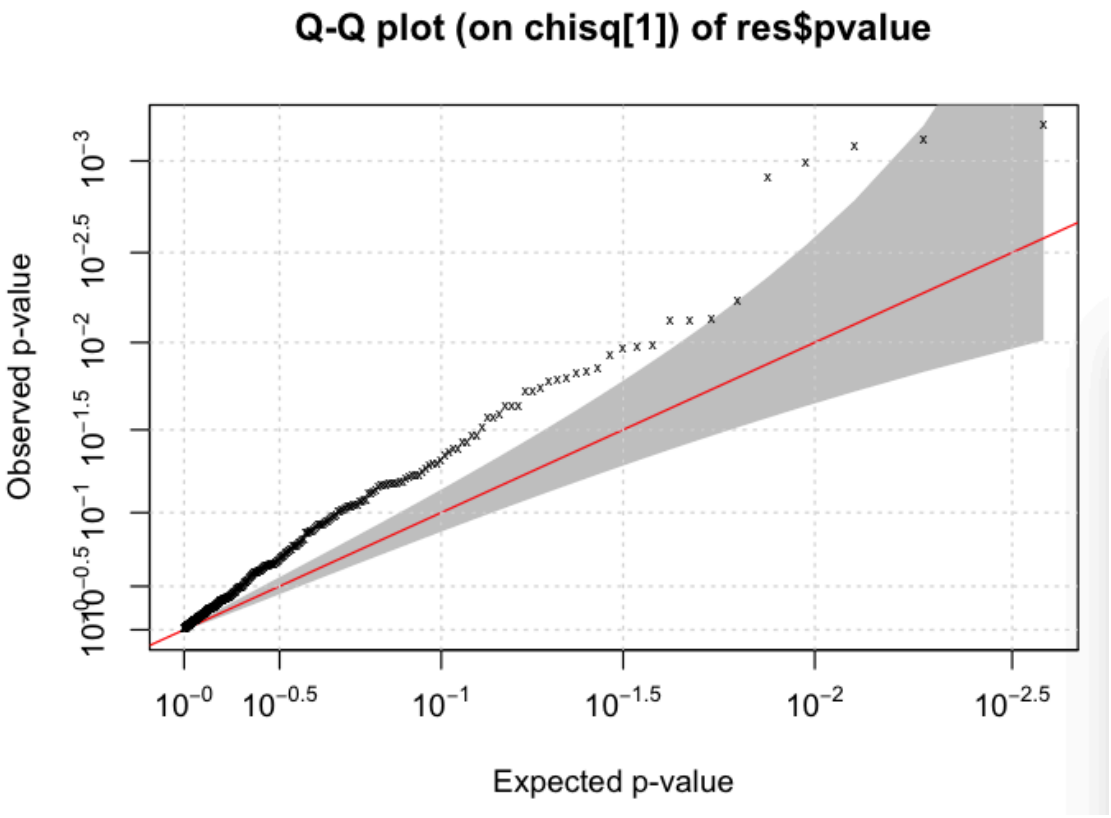
Test 7:

Q-Q plot



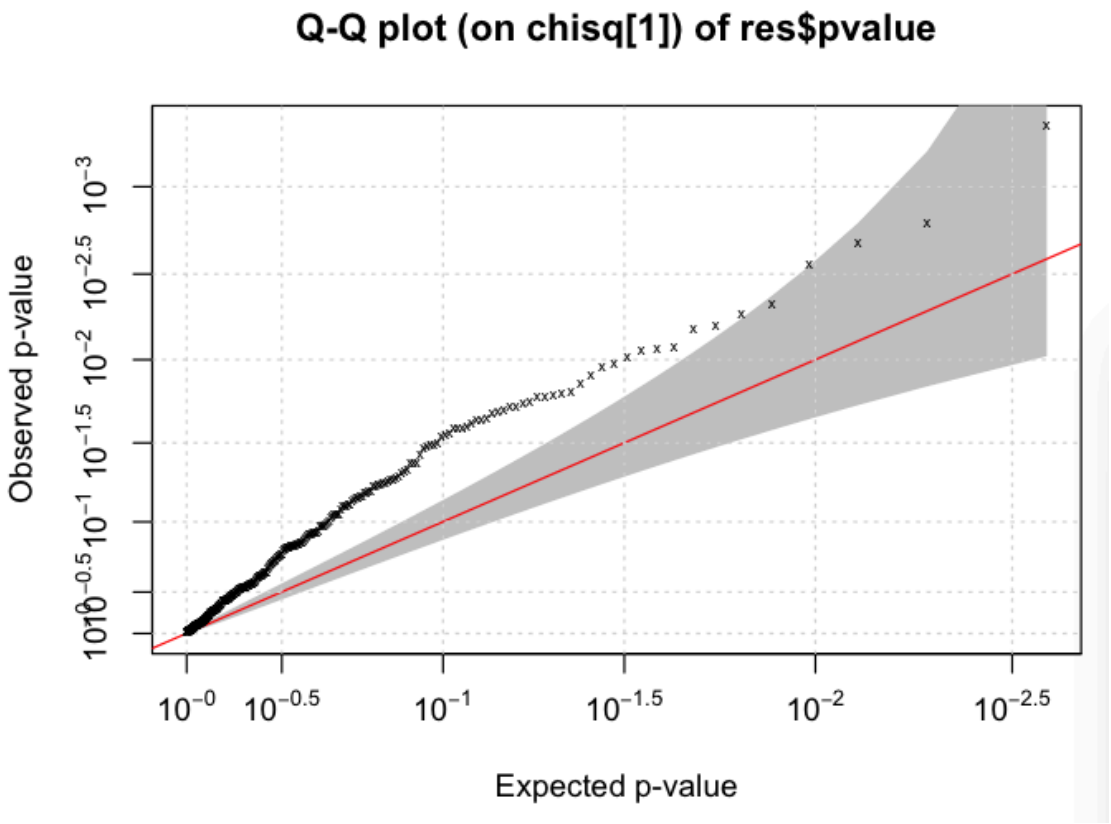
Test 8:

Q-Q plot



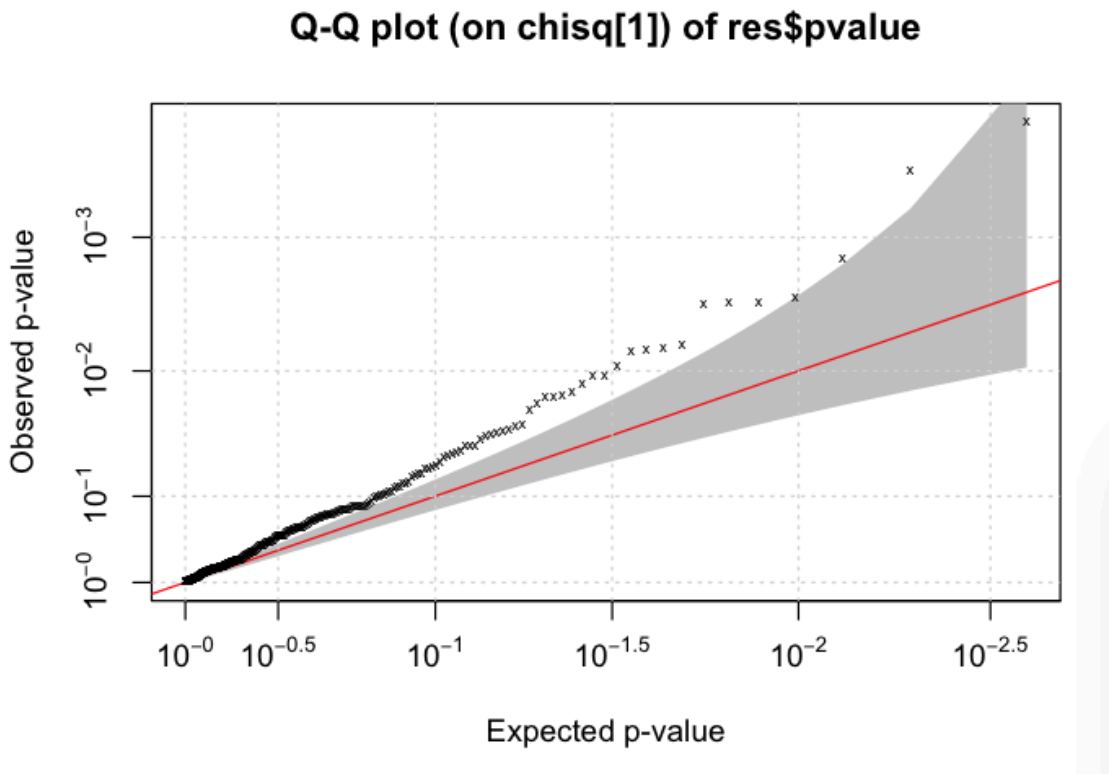
Test 9:

Q-Q plot



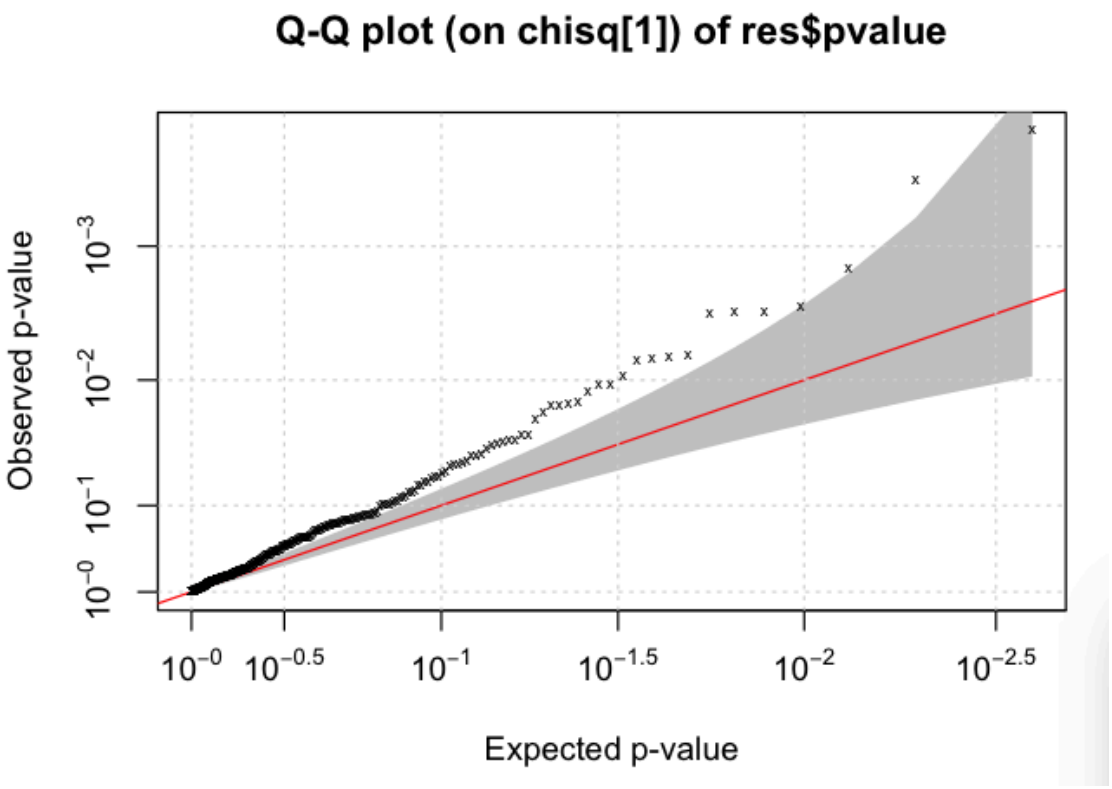
Test 10:

Q-Q plot



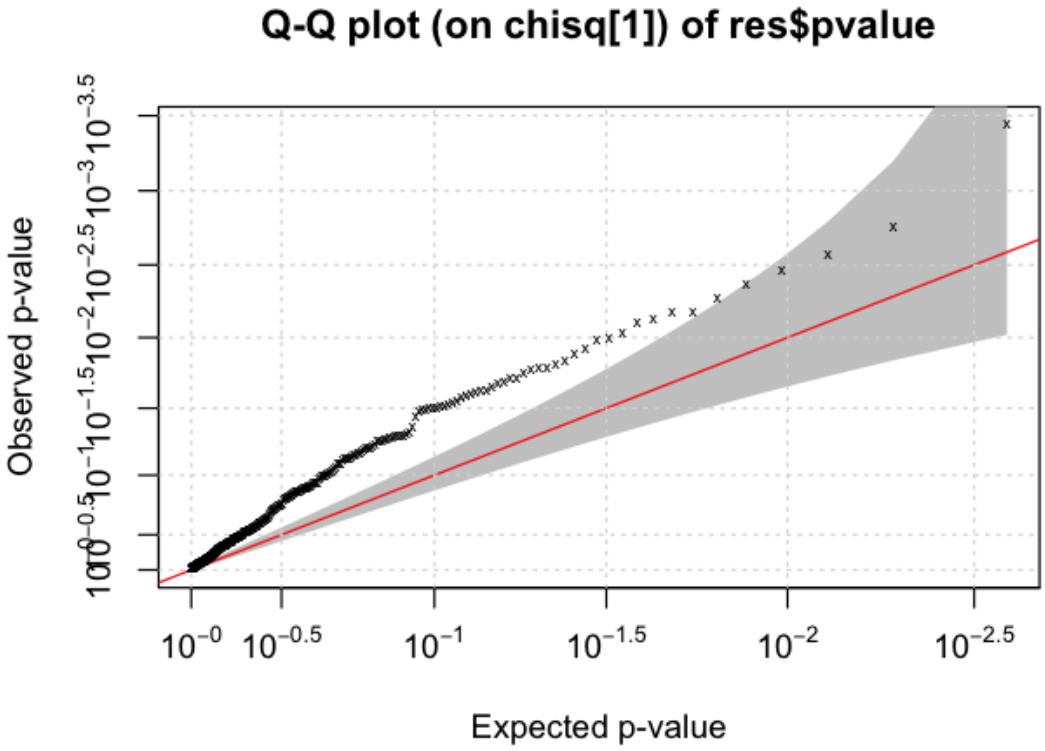
Test 11:

Q-Q plot



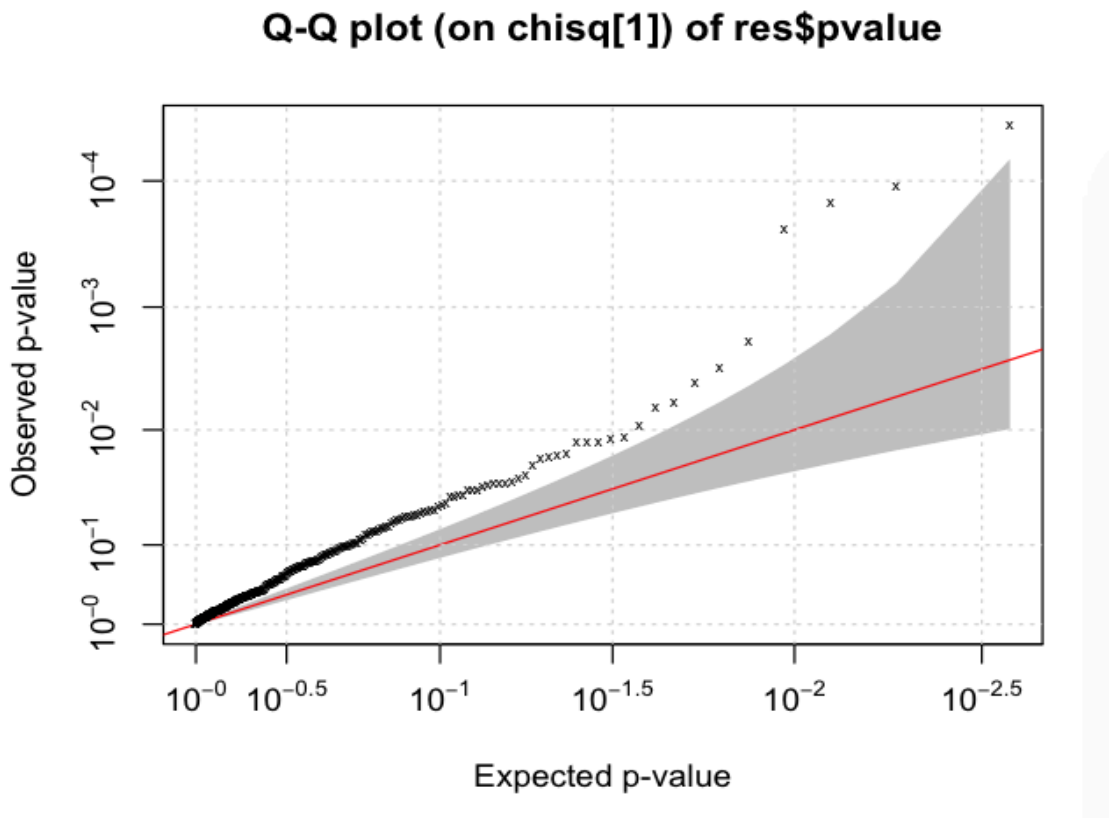
Test 12:

Q-Q plot



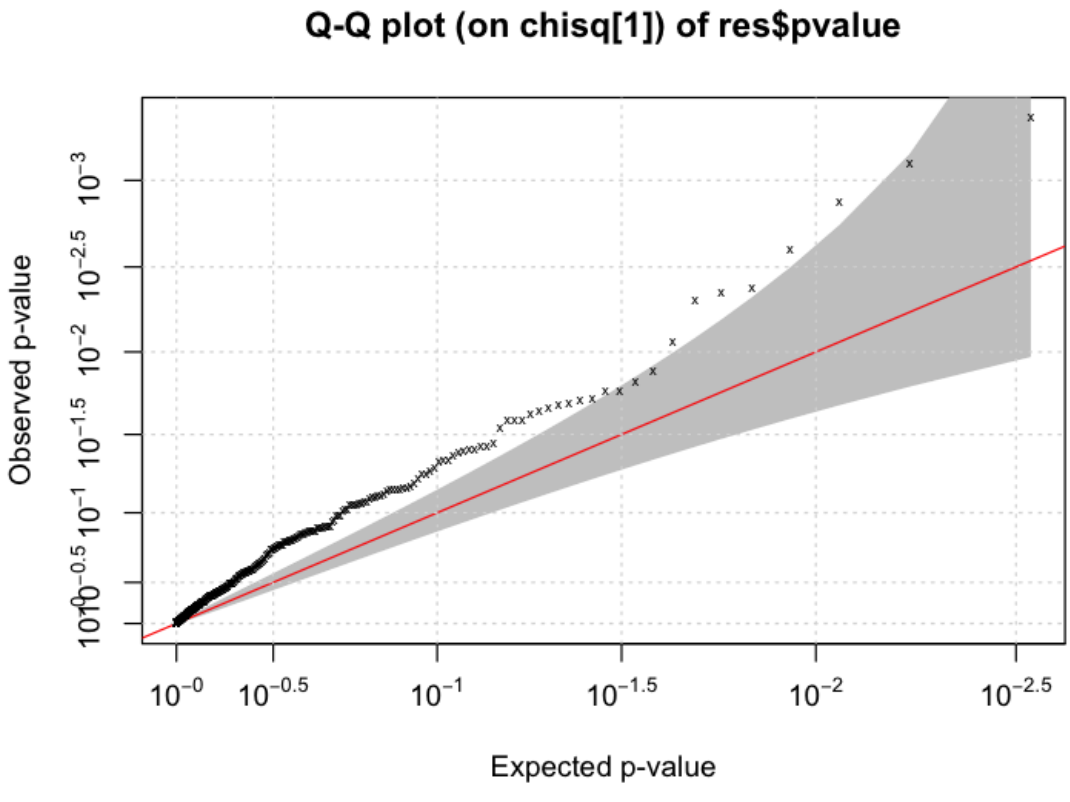
Test 13:

Q-Q plot



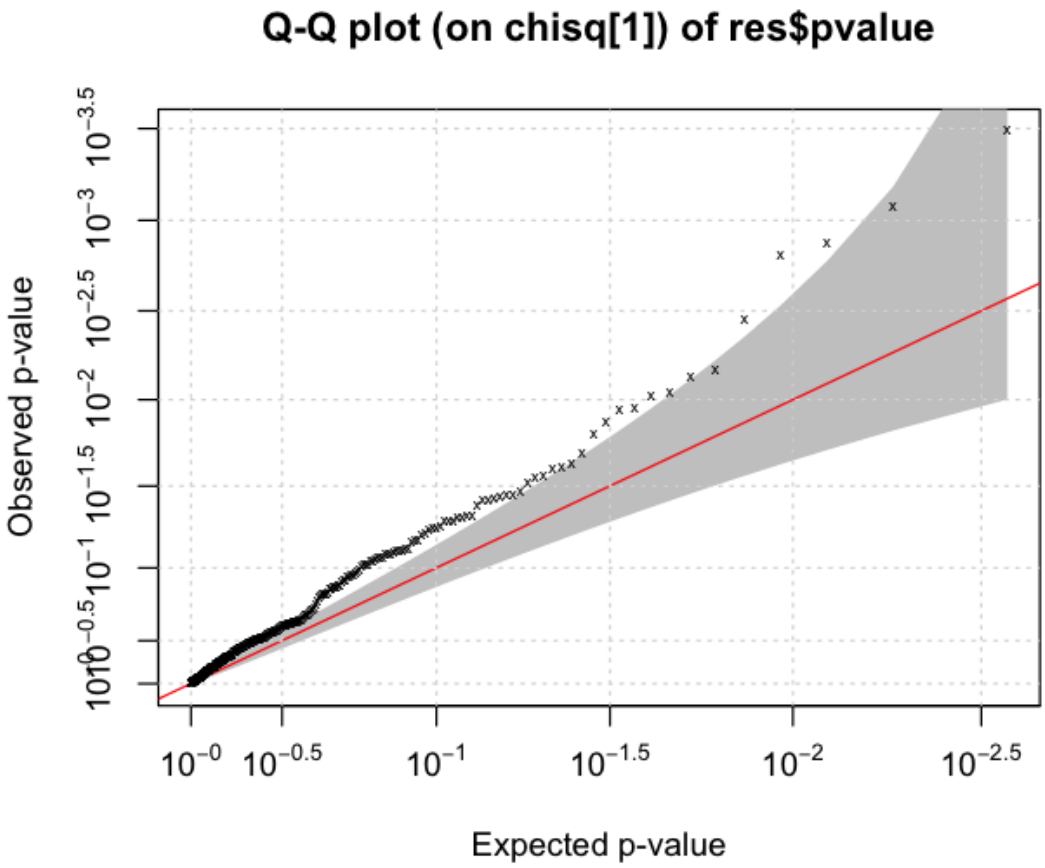
Test 14:

Q-Q plot



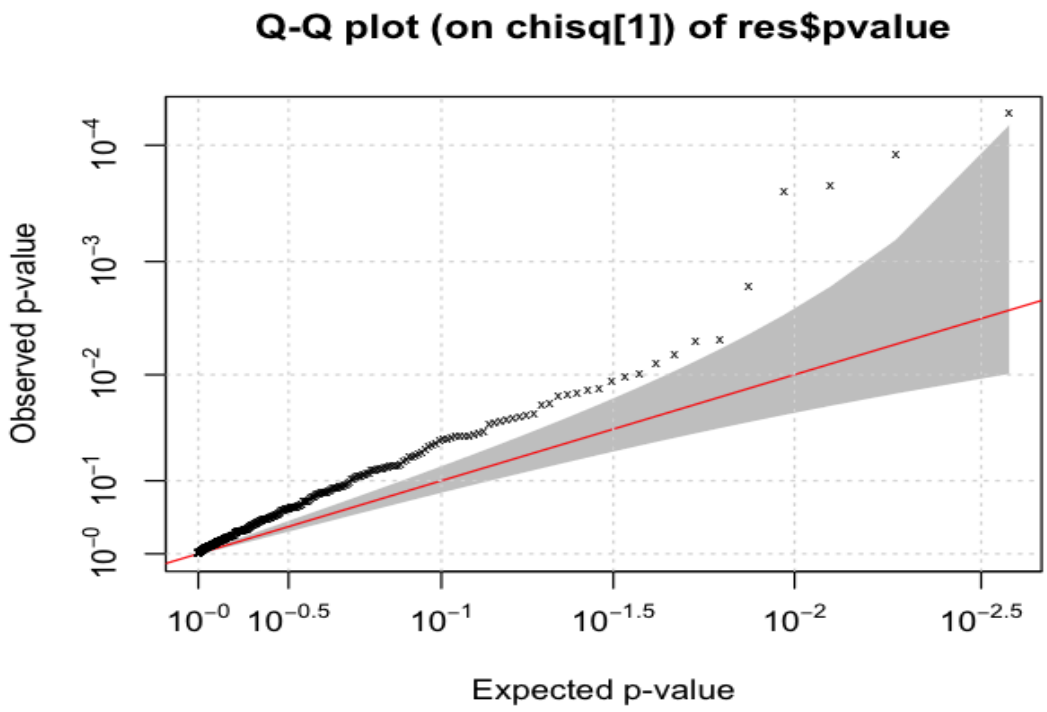
Test 15:

Q-Q plot



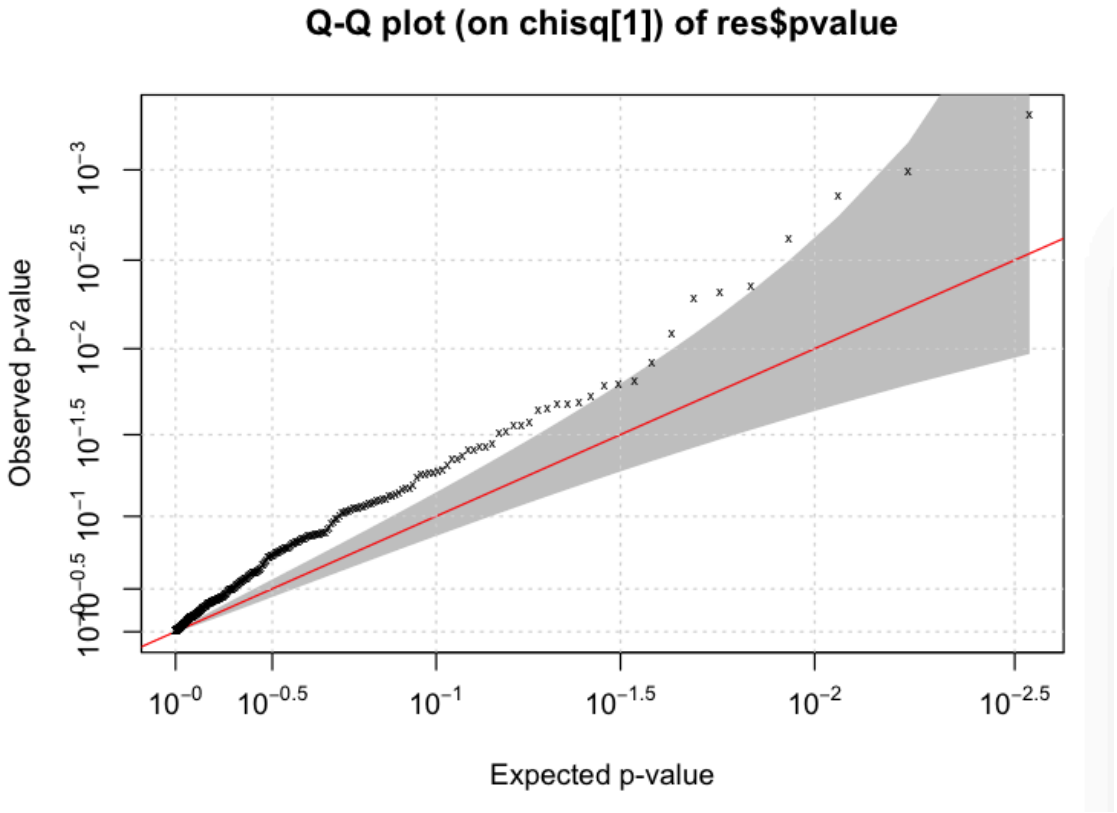
Test 16:

Q-Q plot



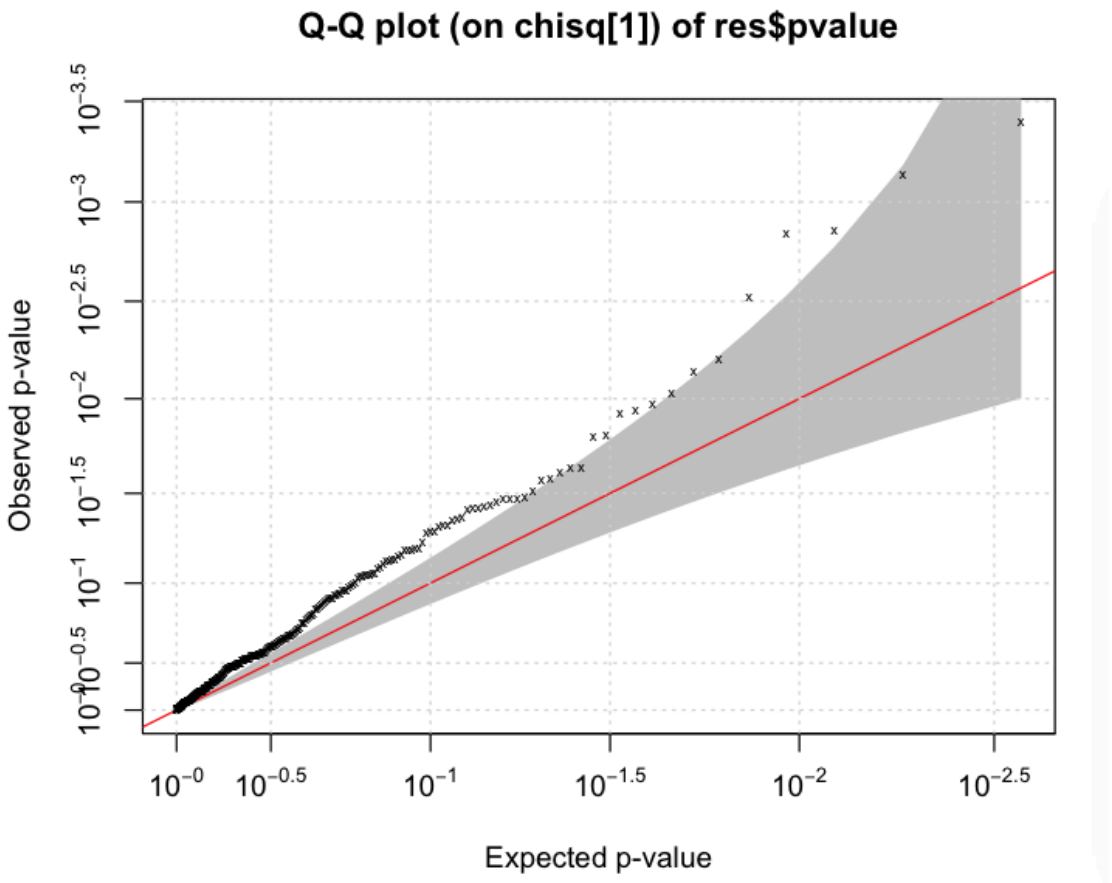
Test 17:

Q-Q plot



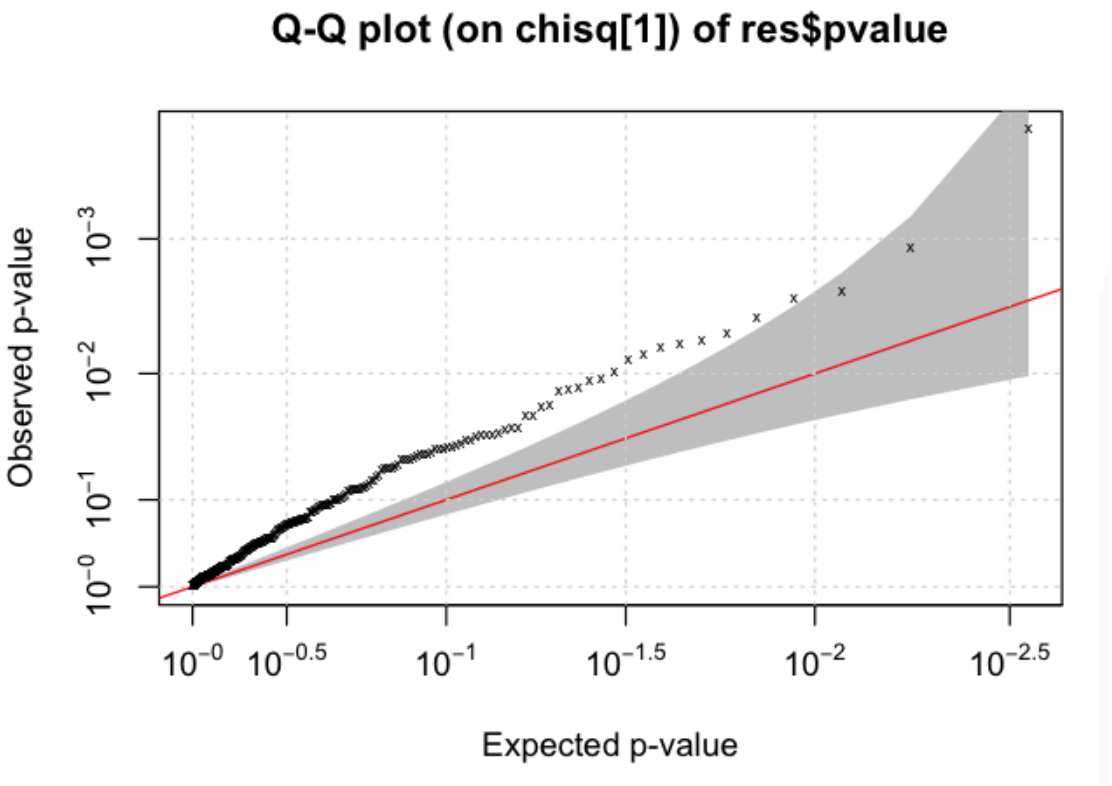
Test 18:

Q-Q plot



Test 19:

Q-Q plot



Appendix J

Guy's and St Thomas' NHS
NHS Foundation Trust

Research & Development
16th Floor Tower Wing
Guy's Hospital
Great Maze Pond
London SE1 9RT
Tel: 020 7188 7188

Dr Maria Rosario Caballero
Consultant
Guy's and St Thomas' NHS Hospital
Guy's Hospital
Department of Allergy
2nd floor, Bermondsey Wing
London, SE19RT

18/08/2011

Dear Dr Caballero

Title: Genetic Epidemiology of Immediate Allergy to Beta-Lactam Antibiotics

In accordance with the Department of Health's Research Governance Framework for Health and Social Care, all research projects taking place within the Trust must receive a favourable opinion from an ethics committee and approval from the Department of Research and Development (R&D) prior to commencement.

- **Ethics Number:** 11/LO/0112
- **Sponsor:** KCL/GSTFT
- **Funder:** MGETSFR
- **End Date:** 08/04/2013
- **Protocol:** Version 2,11/04/2011
- **Site:** GSTFT
- **R&D Approval Date:** 18/08/2011
- **Chief Investigator:** Ms Sevil Badin

NHS permission for the above research has been granted on the basis described in the application form, protocol and supporting documentation as listed in the ethics letter of favourable opinion letter dated 04/05/2011. I am pleased to inform you that we are approving the work to proceed within Guy's and St Thomas' NHS Foundation Trust and that the study has been allocated the Trust R&D registration number RJ111/N253. Please quote the R&D registration number in any communications with the R&D Department regarding your project.

Conditions of Approval:

- The principal investigator must ensure that the recruitment figures are reported.
- The principal investigator must notify R&D of the actual end date of the project.
- R&D must be notified of any changes to the protocol prior to implementation.
- The project must follow the agreed protocol and be conducted in accordance with all Trust Policies and Procedures especially those relating to research and data management.
- Members of the research team must have appropriate substantive or honorary contracts with the Trust prior to the study commencing. Any additional researchers who join the study at a later stage must also hold a suitable contract.

Data Protection:
Please ensure that you are aware of your responsibilities in relation to The Data Protection Act 1998, NHS Confidentiality Code of Practice, NHS Caldicott Report and Caldicott Guardians, the Human Tissue Act 2004, Good Clinical Practice, the NHS Research Governance Framework for Health and Social Care, Second Edition April 2005 and any further legislation released during the time of this study.

The Principal Investigator is responsible for ensuring that Data Protection procedures are observed throughout the course of the project.

If the project is a clinical trial under the European Union Clinical Trials Directive the following must also be complied with:

1. The EU Directive on Clinical Trials (Directive 2001/20/EC) and UK's implementation of the Directive: The Medicines for Human Use (Clinical Trials) Regulations 2004;
2. The EU Directive on Principles and Guidelines for Good Clinical Practice (EU Commission Directive 2005/28/EC); and UK's implementation of the Directive: The Medicines for Human Use (Clinical Trials) Amendment Regulations 2006;
3. If a clinical trials team has to keep a subject in a department "out of hours" for whatever reason, the Senior Nurse for the Hospital should be informed of their presence – as should the Resuscitation Team.

Amendments:
Please ensure that you submit a copy of any amendments made to this study to the R&D Department.

ISRCTN registration:
If appropriate it is recommended that you register with the Current Controlled Trials website http://isrctn.org/. Find out more about registering for an International Standard Randomised Controlled Trial Number (ISRCTN) as part of the Portfolio application process. Non-commercial studies with an interventional component that are eligible for NIHR CRN support can register for an ISRCTN for free via the Portfolio Database.

Annual Progress Report:
It is obligatory that an annual report is submitted by the Chief Investigator to the research ethics committee, and we ask that a copy is sent to the R&D Department. The yearly period commences from the date of receiving a favourable opinion from the ethics committee.

Please submit a copy of the progress report on the anniversary of the Ethics favourable opinion **(4th May)**

Should you require any further information please do not hesitate to contact us.

In line with the Research Governance Framework, your project may be randomly selected for monitoring for compliance against the standards set out in the Framework. For information, the Trust's process for the monitoring of projects and the associated guidance is available from the Trust's intranet or on request from the R&D Department. You will be notified by the R&D Department if and when your project has been selected as part of the monitoring process. No action is needed until that time.

Thank you for registering your research project.

Yours sincerely

Janahi Visakan
R&D Co-ordinator
cc: Keith Brennan
cc: Sevil Badin

Appendix K

Guy's and St Thomas' NHS
NHS Foundation Trust

Dr. Rosario Caballero MD PhD
Guy's Hospital
Dept of Allergy & Respiratory Science
2nd Floor, Thomas Guy House
St Thomas Street
London SE1 9RT
Tel: 020 7188 5075
Fax: 020 7188 5844
Email:
rosario.caballero@kcl.ac.uk

Dr. Kourosh R Ahmadi PhD
St Thomas' Hospital
Dept Twin Research & Genetic Epidemiology
Lambeth Palace Rd
London SE1 7EH
Tel: 020 7188 6728
Email:
kourosh.ahmadi@kcl.ac.uk

CONSENT FORM

LREC Reference number: 104/H07224/48.

Study title: Genetic epidemiology of β-lactam antibiotic allergy

Name of Researcher: Dr Rosario Caballero & Dr Kourosh R Ahmadi

		Please initial box
1.	I confirm that I have read and understand the information sheet for the above study and have had the opportunity to ask questions.	☐
2.	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.	☐
3.	I understand that sections of any of my medical notes may be looked at by responsible individuals from regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records.	☐
4.	I agree that my tissue, including Blood and DNA or RNA, may be stored and used in this study. (The storage will be covered by a Human Tissue Authority licence)	☐
5.	I agree that any samples collected as part of this study may be stored anonymously for use in future, related studies, by the principle investigators.	☐
7.	I agree to take part in the above study.	☐

Name of Patient	Date	Signature
Name of Person taking consent (if different from researcher)	Date	Signature
Researcher	Date	Signature

1 for patient; 1 for researcher; 1 to be kept with hospital notes

Appendix L

Guy's and St Thomas' NHS
NHS Foundation Trust

Dr. Rosario Caballero MD PhD
Guy's Hospital
Dept of Allergy & Respiratory Science
2nd Floor, Thomas Guy House
St Thomas Street
London SE1 9RT
Tel: 020 7188 5075
Fax: 020 7188 5844
Email:
rosario.caballero@kcl.ac.uk

Dr. Kourosh R Ahmadi PhD
St Thomas' Hospital
Dept Twin Research & Genetic Epidemiology
Lambeth Palace Rd
London SE1 7EH
Tel: 020 7188 6728
Email:
kourosh.ahmadi@kcl.ac.uk

PARTICIPANT INFORMATION SHEET

LREC reference: 104/H07224/48

Study title: Genetic epidemiology of β-lactam antibiotic allergy

You are being invited to take part in a research study. Before you decide to take part, it is important for you to understand why the research is being done and what it will involve. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Betalactam antibiotics, including penicillins and cephalosporins, are widely used to treat common infections. Unfortunately, in up to 10% of the people they elicit a marked, unwanted allergic reaction. Such adverse reactions not only represent as a significant burden to the patient but also increase the overall cost of treatment markedly. Recent studies have shown that certain sections of the population face greater risk of suffering such adverse events which can be due to a combination of factors including genetic, environmental exposure (diet, infection, interactions with other drugs), as well as other epidemiological (age, sex) aspects.

The purpose of this study is to uncover genetic and environmental factors that increase the susceptibility to allergy to betalactam antibiotics. More specifically, we will attempt to identify common and rarer genetic variants as well as damning environmental or epidemiological factors that are associated with betalactam allergy. The results should not only increase our understanding of a key biological system but also facilitate clinical prediction of adverse allergic responses to two of the most commonly prescribed drugs in the clinic.

Why have I been invited?

You have been invited to take part in this study because you have a history of immediate reaction following administration of penicillin or cephalosporin antibiotics. If you have no history of immediate reactions with these antibiotics, you have been invited to participate as a control participant.

Do I have to take part?

It is up to you to decide whether or not to take part. If you agree to participate in the study you will be given this information sheet to keep and asked to sign a consent form (attached). If you do decide to take part you are still free to withdraw at any time and without having to give a reason. Whether you decide not to take part or initially agree to take part and later change your

mind, the standard of your care in hospital, either presently or at any time in the future, will not be affected in any way.

What will happen to me if I take part?

If you agree to take part in this study you will be asked to re visit the allergy clinic. At the visit, we will ask you to:

- Sign a consent form;
- Ask any questions you have regarding the study;
- Give a sample of urine and 30 ml of blood (about a tablespoonful) which will be used to extract DNA and RNA or to carry out further biochemical tests deemed useful for characterising the allergy.

How long will the visits last?

The visit should take no longer that 1 hour.

What are the side effects of taking part?

You are not at risk of any side effects during the visit.

What are the possible benefits of taking part?

The results of this study should aid in identifying individuals that face a greater risk of suffering from antibiotics allergy and so inform medical practitioners in their choice of drug prescription and regime.

What if I want to withdraw fro the study?

You may withdraw from the study at any time. You do not have to give a reason. Withdrawing from the study will not influence in any way any current or future treatment you may be receiving, or will receive at the hospital.

Will my participating in this study be kept confidential?

All information which is collected about you during the course of this research will be kept strictly confidential. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised from it.

Storage of your blood sample

As part of the study, urine and blood sample will be collected to enable us to carry out the study. We ask for your permission to store any surplus blood, urine, DNA or RNA in our laboratories for future use in other, similar studies. The samples will be identifiable as having come from you, but will be used only for the purpose of research and not for any diagnostic purpose. No information will be divulged to any third party without your written consent. These samples might be made available, with the permission of Head of Department, to any member of the departmental research team for similar related studies, but not to any third party in another department or to a commercial organisation. Samples may be stored indefinitely. Please note that you can withdraw your consent for us to store your samples at any time, when they would then be destroyed. The storage of your tissues is covered by a departmental licence issued by the Human Tissues Authority.

What will happen to the results of this research study?

Eventually the results of the study will be published in a medical journal. You are quite welcome to receive from us any resulting publications if you so wish, but you should know that your participation would be entirely anonymous.

Who is organising and funding the research?

The study is being funded in part by Biomedical Research Centre maintenance funding. Guy's and St Thomas' NHS Foundation Trust, along with its academic partner King's College London, is one of five new National Institute for Health Research comprehensive Biomedical Research Centres in the UK. These centres have a strong focus on "translational research" taking advances in basic medical research out of the laboratory and into the clinical setting so that they can benefit patients. They form a key part of the Department of Health's new strategy for research and development in the NHS.

Who has reviewed the study?

The study was reviewed and approved by the North West London Research Ethics Committee 1.

Contact for further information

For further information about this study, please contact Dr Kourosh Ahmadi (email: kourosh.ahmadi@kcl.ac.uk or telephone on 020 7188 6728) or Dr Rosario Caballero (rosario.caballero@kcl.ac.uk or telephone on 020 7188 5075).

Thank you for your time in reading this information sheet and considering participating in this research study.

Please keep this information sheet and a copy of your consent form for future reference.

Appendix M

Allergy Questionnaire

Name: NHS #: Date: ___ / ___ / _____ DOB: ___ / ___ / ____

Age: _____ Sex: ☐ Male ☐ Female Occupation: _____

Race: ☐ White ☐ Black ☐ Asian ☐ Mixed ☐ Other

Existing Medical Conditions:

- ☐ Cancer
- ☐ Cardiovascular Disease
- ☐ High Blood Pressure
- ☐ Alcohol/Drug Abuse
- ☐ High Cholesterol
- ☐ Lung/Respiratory Disease
- ☐ Infectious Disease
- ☐ Asthma
- ☐ Immune Disorders
- ☐ Obesity
- ☐ Other
- ☐ Stroke
- ☐ Diabetes
- ☐ Depression
- ☐ Liver Disease
- ☐ Kidney Disease
- ☐ Neurological Disorders
- ☐ Eczema
- ☐ Menopause
- ☐ Hay Fever
- ☐ Contact Dermatitis
- ☐ Urticaria

If you ticked yes Please expand: _____

_____

Current Medication:

- ☐ Vitamins/Minerals
- ☐ NSAIDs
- ☐ Asthma Medication
- ☐ Oral Contraceptives
- ☐ Sedatives/Sleep Aids
- ☐ Prescription Pain Medication
- ☐ Oral Hypoglycemics
- ☐ Hormones
- ☐ Diuretics
- ☐ Statins
- ☐ Other
- ☐ Herbs
- ☐ Aspirin
- ☐ Antihistamines
- ☐ Thyroxin
- ☐ Steroids (Nasal/Topical)
- ☐ Antidepressants
- ☐ Insulin
- ☐ Antibiotics
- ☐ Other Blood Pressure Tablets
- ☐ Anticoagulants
- ☐ Antifungals

If you ticked yes Please expand: _____

_____

Existing Medical Conditions: ☐ Eczema ☐ Asthma ☐ Hayfever ☐ Urticaria

Food Allergy: ☐ Suspected ☐ Known _____

Known Allergies: ☐ Nickel/Metals ☐ Flowers/Trees/Grasses ☐ Fragrance ☐ Latex ☐ Rubber ☐ Medicines ☐ Insects ☐ Animals

☐ Other _____

☐ Suspected Allergies _____

Serious Drug Reactions : ☐ No ☐ Yes: Drugs/Date _____

Family History:

- Allergies and Asthma: ☐ Yes ☐ No
- Eczema: ☐ Yes ☐ No
- Hay Fever: ☐ Yes ☐ No
- Antibiotics Allergy ☐ Yes ☐ No